6 Laplace transform

Summary

In order to consider the response of engineering systems, e.g. electrical or control systems, to inputs such as step, or perhaps an impulse, we need to be able to solve the differential equation for that system with that particular form of input. As the previous chapter indicates, this can be rather laborious. A simpler method of tackling the solution is to transform a differential equation into a simple algebraic equation which we can easily solve. This is achieved by the use of the Laplace transform, the subject of this chapter.

Objectives

By the end of this chapter, the reader should be able to:

- understand what using the Laplace transform involves;
- use Laplace transform tables to convert first- and second-order differential equations into algebraic equations;
- use Laplace transform tables, and where appropriate partial fractions, to convert Laplace transform equations into real world equations;
- determine the outputs of systems to standard input signals such as step, impulse and ramp.

6.1 The Laplace transform

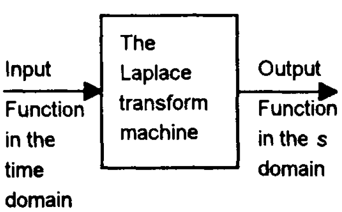


Figure 6.1 *The Laplace transform*

In this chapter a method of solving such differential equations is introduced which transforms a differential equation into an algebraic equation. This is termed the *Laplace transform*. It is widely used in engineering, in particular in control engineering and in electrical circuit analysis where it is commonplace not even to write differential equations to describe conditions but to write directly in terms of the Laplace transform.

We can think of the Laplace transform as being rather like a function machine (Figure 6.1). As input to the machine we have some function of time $f(t)$ and as output a function we represent as $F(s)$. The input is referred to as being the *time domain* while the output is said to be in the *s-domain*. Thus we take information about a system in the time domain and use our 'machine' to transform it into information in the s-domain. Differential equations which describe the behaviour of a system in the time domain are converted into algebraic equations in the s-domain, so considerably simplifying their solution. We can thus transform a

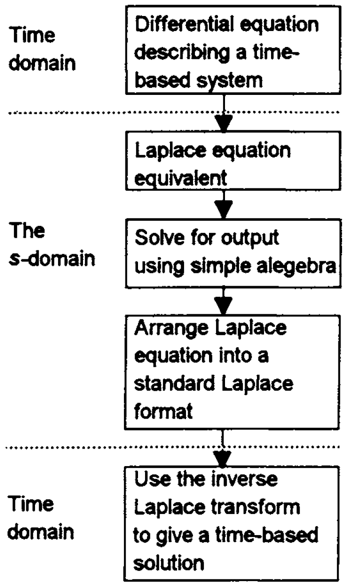


Figure 6.2 *Using the Laplace transform. As an illustration, Ohm's Law gives the time-domain equation v(t) = Ri(t), both v and i being functions of time and R assumed to remain constant. In the s-domain this becomes V(s) = RI(s). After working with this equation in the s-domain we can then transform back to the time domain*

Key point

To obtain the Laplace transform of a function of time $f(t)$, multiply it by e^{-st} and integrate the product between zero and infinity.

Key point

When electrical circuits are discussed in terms of currents or voltages varying with time we use differential equations and are said to be working in the *time domain*. When we use phasors we can be said to be working in the *frequency domain*; we are no longer working with time-varying quanitities. As we shall see later in this chapter, we can transform currents or voltages which vary with time into the *s-domain* by using the Laplace transform; like the transformation using phasors, we are no longer working with time-varying quantities.

differential equation into an s-domain equation, solve the equation and then use the 'machine' in inverse operation to transform the s-domain equation back into a time-domain solution (Figure 6.2).

Now this definition may look rather daunting, but do not fear, it is very likely that you will not need to use it but rather will make use of tables which other people have worked out. However, you should appreciate the basis of the transform. The *Laplace transform* of some function of time is defined by:

> *Multiply a given function of time $f(t)$ by e^{-st} and integrate the product between zero and infinity. The result, if it exists, is called the Laplace transform of $f(t)$ and is denoted by $\mathcal{L}\{f(t)\} = F(s)$.*

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)\,dt \qquad [1]$$

Note that the integration is between 0 and $+\infty$ and so is *one-sided* and not over the full range of time from $-\infty$ to $+\infty$.

Example

Determine the Laplace transform of $f(t) = 1$.

Using equation [1]:

$$\mathcal{L}\{f(t)\} = \int_0^\infty 1\ e^{-st}\,dt = \left[\frac{e^{-st}}{-s}\right]_0^\infty = \frac{1}{s}$$

This is provided that $s > 0$ so that $e^{-st} \to 0$ as $t \to \infty$

Example

Determine the Laplace transform of $f(t) = e^{at}$.

Using equation [1]:

$$\mathcal{L}\{f(t)\} = \int_0^\infty e^{at}\,e^{-st}\,dt = \int_0^\infty e^{-(s-a)t}\,dt = \left[\frac{e^{-(s-a)t}}{-(s-a)}\right]_0^\infty = \frac{1}{s-a}$$

That is provided we have $(s - a) > 0$.

Example

Determine the Laplace transform of $f(t) = t$.

Using equation [1]:

$$\mathcal{L}\{f(t)\} = \int_0^\infty t\ e^{-st}\,dt$$

Using integration by parts:

$$\mathcal{L}\{f(t)\} = \left[-\frac{t}{s}\,e^{-st}\right]_0^\infty + \int_0^\infty \frac{1}{s}\,e^{-st}\,dt = \left[-\frac{1}{s^2}\,e^{-st}\right]_0^\infty = \frac{1}{s^2}$$

That is provided we have $s > 0$.

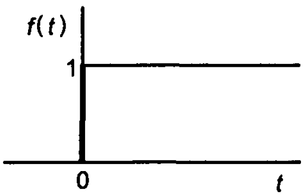


Figure 6.3 *Unit step at time t = 0*

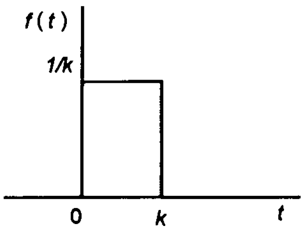


Figure 6.4 *Unit area rectangular pulse*

Laplace transforms for step and impulse function

Consider the *unit step function* $u(t)$ shown in Figure 6.3. The Laplace transform is given by equation [1] as:

$$\mathcal{L}\{u(t)\} = \int_0^\infty 1\,e^{-st}\,dt = \left[\frac{e^{-st}}{-s}\right]_0^\infty = \frac{1}{s} \qquad [2]$$

Thus a unit size step input signal to an engineering system occurring at time $t = 0$ will have a Laplace transform of $1/s$.

Now consider obtaining the *unit impulse function* (represented as $\delta(t)$). Such an impulse can be considered to be a unit area rectangular pulse which has its width k decreased to give the unit impulse in the limit when $k \to 0$. For the unit area rectangular pulse shown in Figure 6.4, the Laplace transform is:

$$\mathcal{L}\{\text{unit area pulse}\} = \int_0^\infty f(t)\,e^{-st}\,dt = \int_0^k \frac{1}{k}\,e^{-st}\,dt + \int_k^\infty 0\,e^{-st}\,dt$$

$$= \left[-\frac{1}{sk}\,e^{-st}\right]_0^k = -\frac{1}{sk}(e^{-sk} - 1)$$

We can replace the exponential by a series, thus obtaining:

$$\mathcal{L}\{\text{unit area pulse}\} = -\frac{1}{sk}\left(1 + (-sk) + \frac{(-sk)^2}{2!} + \frac{(-sk)^3}{3!} + \ldots - 1\right)$$

Thus in the limit as $k \to 0$, the Laplace transform tends to the value 1 and so:

$$\mathcal{L}\{\delta(t)\} = 1 \qquad [3]$$

Thus a unit size impulse input signal occurring at time $t = 0$ to an engineering system will have a Laplace transform of 1.

Key points

The Laplace transforms of signals commonly used as inputs to systems are:

Unit impulse: 1
Unit step: $1/s$
Unit ramp: $1/s^2$

Standard Laplace transforms

The transforms derived above, together with others, are tabulated as a set of standard transforms so that it becomes unnecessary to derive them by the use of equation [1]. Table 6.1 gives some of the more common standard transforms. As indicated in the following section, these standard transforms can be used to derive the transforms for a wide range of functions.

Table 6.1 *Laplace transforms*

	$f(t)$	$\mathcal{L}\{f(t)\}$
1	Unit impulse $\delta(t)$	1
2	Unit step $u(t)$	$\frac{1}{s}$
3	Unit ramp t	$\frac{1}{s^2}$
4	t^n	$\frac{n!}{s^{n+1}}$
5	e^{-at}	$\frac{1}{s+a}$
6	$1 - e^{-at}$	$\frac{a}{s(s+a)}$
7	$t\,e^{-at}$	$\frac{1}{(s+a)^2}$
8	$e^{-at} - e^{-bt}$	$\frac{b-a}{(s+a)(s+b)}$
9	$(1-at)\,e^{-at}$	$\frac{s}{(s+a)^2}$
10	$1 - \frac{b}{b-a}\,e^{-at} + \frac{a}{b-a}\,e^{-bt}$	$\frac{ab}{s(s+a)(s+b)}$
11	$\frac{e^{-at}}{(b-a)(c-a)} + \frac{e^{-bt}}{(c-a)(a-b)} + \frac{e^{-ct}}{(a-c)(b-c)}$	$\frac{1}{(s+a)(s+b)(s+c)}$
12	$\sin \omega t$	$\frac{\omega}{s^2+\omega^2}$
13	$\cos \omega t$	$\frac{s}{s^2+\omega^2}$
14	$e^{-at} \sin \omega t$	$\frac{\omega}{(s+a)^2+\omega^2}$
15	$e^{-at} \cos \omega t$	$\frac{s+a}{(s+a)^2+\omega^2}$
16	$\sinh \omega t$	$\frac{\omega}{s^2-\omega 2}$
17	$\cosh \omega t$	$\frac{s}{s^2-\omega^2}$
18	$\frac{\omega}{\sqrt{1-\zeta^2}}\,e^{-\zeta\omega t} \sin \omega\sqrt{1-\zeta^2}\,t,\ \zeta < 1$	$\frac{\omega^2}{s^2+2\zeta\omega s+\omega^2}$
19	$1 - \frac{1}{\sqrt{1-\zeta^2}}\,e^{-\zeta\omega t} \sin\left(\omega\sqrt{1-\zeta^2}\,t+\phi\right), \cos\phi = \zeta$	$\frac{\omega^2}{s(s^2+2\zeta\omega s+\omega^2)}$

Properties of Laplace transforms

The following are basic properties of Laplace transforms and can be used with the above table of standard transforms to obtain a wide range of other transforms.

Key point

If two separate time functions have Laplace transforms then the transform of the sum of the time functions is the sum of the Laplace transforms of the two functions considered separately.

- ***Sum of two functions***

  If two separate time functions $f(t)$ and $g(t)$ have Laplace transforms then the transform of the sum of the time functions, i.e. $f(t) + g(t)$, is the sum of the Laplace transforms of the two functions considered separately:

$$\mathcal{L}\{f(t) + g(t)\} = \mathcal{L}\{f(t)\} + \mathcal{L}\{g(t)\} \qquad [4]$$

This property is derived by using equation [1]:

$$\mathcal{L}\{f(t)+g(t)\} = \int_0^\infty \{f(t)+g(t)\}\,\mathrm{e}^{-st}\,\mathrm{d}t$$

$$= \int_0^\infty f(t)\,\mathrm{e}^{-st}\,\mathrm{d}t + \int_0^\infty g(t)\,\mathrm{e}^{-st}\,\mathrm{d}t$$

$$= \mathcal{L}\{f(t)\} + \mathcal{L}\{g(t)\}$$

Since $2f(t)$ equals $f(t) + f(t)$, then the Laplace transform of $2f(t)$ will be twice the Laplace transform of $f(t)$. Thus, in general:

$$\mathcal{L}\{af(t)\} = a\mathcal{L}\{f(t)\} \qquad [5]$$

The Laplace transform of a constant multiplying a function is the same as a the constant multiplying the Laplace transform of the function.

Key point

The Laplace transform of a constant multiplying a function is the same as a the constant multiplying the Laplace transform of the function.

Example

Determine the Laplace transform of $1 + 2t$.

Using equations [4] and [5] and Table 6.1:

$$\mathcal{L}\{1+2t\} = \frac{1}{s} + \frac{2}{s^2}$$

Example

Determine the Laplace transform of $3\sin 2t + \cos 2t$.

Using equations [4] and [5] and Table 6.1:

$$\mathcal{L}\{3\sin 2t + \cos 2t\} = 3\frac{2}{s^2+4} + \frac{s}{s^2+4}$$

Example

Determine the Laplace transform of $3t^2 + 2\,\mathrm{e}^{-t}$.

Using equations [4] and [5] and Table 6.1:

$$\mathcal{L}\{3t^2 + 2\,\mathrm{e}^{-t}\} = 3\frac{2!}{s^3} + 2\frac{1}{s+1}$$

Example

Determine the Laplace transform of $\sin(\omega t + \theta)$.

We can write $\sin(\omega t + \theta)$ as $\sin \omega t \cos \theta + \cos \omega t \sin \theta$. Thus, using equations [4] and [5] and Table 6.1:

$$\mathcal{L}\{\sin(\omega t + \theta)\} = \frac{\omega}{s^2+\omega^2}\cos\theta + \frac{s}{s^2+\omega^2}\sin\theta$$

Example

What is the Laplace transform of an alternating voltage which is described by $240 \sin 314.16t$?

We have an equation of the form constant multiplied by a sine function. Equation [12] in Table 6.1 gives for $\sin \omega t$ the transform $\omega/(s^2 + \omega^2)$. Hence:

$$\mathcal{L}\{240 \sin 314.16t\} = 240\frac{314.16}{s^2+314.16^2}$$

$$= \frac{75.4\times 10^3}{s^2+98.7\times 10^3}$$

- ***The first shift theorem, factor*** $\mathbf{e}^{-at}$

This theorem states that if $\mathcal{L}\{f(t)\} = F(s)$ then:

$$\mathcal{L}\{e^{-at}f(t)\} = F(s + a) \qquad [6]$$

Thus the substitution of $s + a$ for s corresponds to multiplying a time function by e^{-at}. This can be demonstrated if we consider equation [1] with such a function:

$$\mathcal{L}\{e^{-at}f(t)\} = \int_0^\infty e^{-at}f(t)\,e^{-st}\,dt = \int_0^\infty f(t)\,e^{-(s+a)}\,dt$$

Key point

The first shift theorem: if $\mathcal{L}\{f(t)\} = F(s)$ then:

$\mathcal{L}\{e^{-at}f(t)\} = F(s + a)$

Example

Determine the Laplace transform of $e^{-2t} \cosh 3t$.

Using the first shift theorem and the transform for $\cosh 3t$ given by Table 39.1, the transform is that of $\cosh 3t$ with the s replaced by $s + 2$:

$$\mathcal{L}\{e^{-2t} \cosh 3t\} = \frac{s+2}{(s+2)^2-9}$$

Example

Determine the Laplace transform of $2\,e^{-2t}\sin^2 t$.

Since $\cos 2t = 1 - 2\sin^2 t$ we have:

$$\mathcal{L}\{2\sin^2 t\} = \mathcal{L}\{1\} - \mathcal{L}\{\cos 2t\} = \frac{1}{s} - \frac{s}{s^2+4}$$

Hence, using the first shift theorem and replacing the s by $s + 2$:

$$\mathcal{L}\{2\,e^{-2t}\sin^2 t\} = \frac{1}{s+2} - \frac{s+2}{(s+2)^2+4}$$

Key point

The second shift theorem: if a signal is delayed by a time T then its Laplace transform is multiplied by e^{-sT}.

- ***The second shift theorem, time shifting***

The second shift theorem states that if a signal is delayed by a time T then its Laplace transform is multiplied by e^{-sT}. A function $u(t)$ which is delayed is represented by $u(t - T)$, where T is the delay. Thus if $F(s)$ is the Laplace transform of $f(t)$ then:

$$\mathcal{L}\{f(t-T)u(t-T)\} = e^{-sT}F(s) \qquad [7]$$

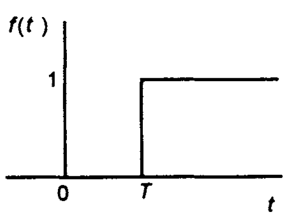


Figure 6.5 *Delayed unit step*

This can be demonstrated by considering a unit step function which is delayed by a time T (Figure 6.5). Equation [1] gives for such a function:

$$\mathcal{L}\{u(t-T)\} = \int_0^\infty u(t-T)\,e^{-st}\,dt = \int_0^T 0\,dt + \int_T^\infty 1\,e^{-st}\,dt$$

$$= \left[\frac{e^{-st}}{-s}\right]_T^\infty = e^{-sT}\frac{1}{s}$$

Example

Determine the Laplace transform for a unit impulse which occurs at a time of $t = 2$ s.

The Laplace transform for a unit impulse at $t = 0$ is 1. Thus the transform for the delayed impulse is $1\,e^{-2s}$.

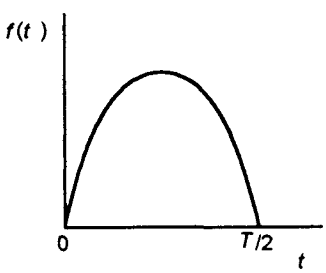


Figure 6.6 *Example*

Example

Determine the Laplace transform of a single pulse consisting of just the first half of a sine wave (Figure 6.6).

We can think of such a function as being the sum of a sine function extending over an infinite number of cycles and a sine function that has had its start delayed by ½T. In this way all but the first half period waveform are cancelled out. Thus the Laplace transform is:

$$\frac{\omega}{s^2+\omega^2}+e^{-sT/2}\frac{\omega}{s^2+\omega^2}$$

- ***Periodic functions***
  A periodic function of period T has a Laplace transform of:

$$\frac{1}{1-e^{-sT}}F_1(s) \qquad [8]$$

where $F_1(s)$ is the Laplace transform of the function for the first period. This can be proved by considering the periodic function to be the sum of the function $f_1(t)$ describing the first period, the first period function delayed by 1 period, the first period function delayed by 2 periods, etc. The Laplace transform of the sum is thus:

$$F_1(s)+e^{-sT}F_1(s)+e^{-2sT}F_1(s)+\ldots=(1+e^{-sT}+e^{-2sT}+\ldots)F_1(s)$$

The term in the brackets is a geometric series with the sum to infinity of $1/(1-e^{-sT})$. Thus we obtain the equation given above.

Key point

A periodic function of period T has a Laplace transform of:

$$\frac{1}{1-e^{-sT}}F_1(s)$$

where $F_1(s)$ is the Laplace transform of the function for the first period.

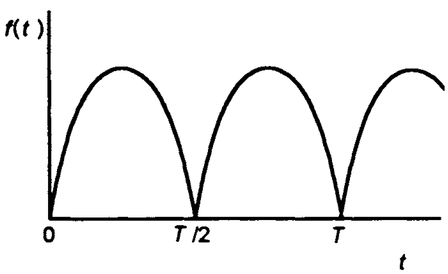


Figure 6.7 *Example*

Example

Determine the Laplace transform of a full-wave rectified sine wave (Figure 6.7).

Such a wave consists of a sequence of the pulses shown in Figure 6.6. Thus the first period function has the transform:

$$\frac{\omega}{s^2+\omega^2}+e^{-sT/2}\frac{\omega}{s^2+\omega^2}$$

Therefore the periodic wave has the Laplace transform:

$$\frac{1}{1-e^{-sT/2}}\left(\frac{\omega}{s^2+\omega^2}+e^{-sT/2}\frac{\omega}{s^2+\omega^2}\right)$$

- ***The Laplace transforms of derivatives***
  Consider the determination of the Laplace transform of the derivative of a function, i.e. $\mathcal{L}\{df(t)/dt\}$. Using equation [1]:

$$\mathcal{L}\left\{\frac{d}{dt}f(t)\right\} = \int_0^\infty e^{-st}\frac{d}{dt}f(t)\,dt$$

Using integration by parts:

$$\mathcal{L}\left\{\frac{d}{dt}f(t)\right\} = -f(0) + s\int_0^\infty e^{-st}f(t)\,dt = -f(0) + sF(s) \qquad [9]$$

where $f(0)$ is the value of $f(t)$ when $t = 0$ and $F(s)$ is the Laplace transform of $f(t)$.

For a second derivative we can similarly obtain:

$$\mathcal{L}\left\{\frac{d^2}{dt^2}f(t)\right\} = \int_0^\infty e^{-st}\frac{d^2}{dt^2}f(t)\,dt$$

$$= \left[e^{-st}\frac{d}{dt}f(t)\right]_0^\infty + s\int_0^\infty e^{-st}\frac{d}{dt}f(t)\,d$$

$$= -\frac{d}{dt}f(0) + s\{-f(0) + sF(s)\}$$

$$= s^2F(s) - sf(0) - \frac{d}{dt}f(0) \qquad [10]$$

where $df(0)/dt$ is the value of the first derivative when $t = 0$.

Likewise for a third derivative we can obtain:

$$\mathcal{L}\left\{\frac{d^3}{dt^3}f(t)\right\} = s^3F(s) - s^2f(0) - s\frac{d}{dt}f(0) - \frac{d^2}{dt^2}f(0) \qquad [11]$$

where $d^2f(0)/dt^2$ is the value of the second derivative at $t = 0$.

Key points

$$\mathcal{L}\left\{\frac{d}{dt}f(t)\right\} = -f(0) + sF(s)$$

where $f(0)$ is the value of $f(t)$ when $t = 0$ and $F(s)$ is the Laplace transform of $f(t)$.

$$\mathcal{L}\left\{\frac{d^2}{dt^2}f(t)\right\} = s^2F(s) - sf(0) - \frac{d}{dt}f(0)$$

where $df(0)/dt$ is the value of the first derivative when $t = 0$.

Example

Given the initial condition that $x = 2$ when $t = 0$, determine the Laplace transform of $4\frac{dx}{dt}$.

Using equation [9]:

$$\mathcal{L}\left\{4\frac{dx}{dt}\right\} = 4[sX(s) - x(0)] = sX(s) - 2$$

where $X(s)$ is the Laplace transform of $x(t)$.

Example

Given the initial conditions that $x = 0$ and $dx/dt = 0$ when $t = 0$, determine the Laplace transform of $3\frac{d^2x}{dt^2} + 2\frac{dx}{dt}$.

Using equations [9] and [10]:

$$\mathcal{L}\left\{3\frac{d^2x}{dt^2}+2\frac{dx}{dt}\right\}$$

$$=3\left[s^2X(s)-sx(0)-\frac{d}{dt}x(0)\right]+2\left[sX(s)-x(0)\right]$$

$$=(3s^2+2s)X(s)$$

- ***Laplace transform of an integral***
Consider the determination of the Laplace transform of the integral of a function, i.e.

$$\mathcal{L}\left\{\int_0^t f(t)\,dt\right\}$$

If we let $g(t) = \int_0^t f(t)\,dt$, then $\frac{d}{dt}g(t) = f(t)$. Then, using equation [7]:

$$\mathcal{L}\left\{\frac{d}{dt}g(t)\right\} = sG(s) - g(0)$$

Since $g(0) = 0$ and $G(s) = \mathcal{L}\{g(t)\}$:

$$\mathcal{L}\{f(t)\} = s\,\mathcal{L}\left\{\int_0^t f(t)\,dt\right\}$$

Thus:

$$\mathcal{L}\left\{\int_0^t f(t)\,dt\right\} = \frac{1}{s}F(s) \qquad [12]$$

Key point

$$\mathcal{L}\left\{\int_0^t f(t)\,dt\right\} = \frac{1}{s}F(s)$$

Example

Determine the Laplace transform of $\int_0^t e^{-t}\,dt$.

Using equation [12]:

$$\mathcal{L}\left\{\int_0^t e^{-t}\,dt\right\} = \frac{1}{s}F(s)$$

Since:

$$F(s) = \mathcal{L}\{e^{-t}\} = \frac{1}{s+1}$$

$$\mathcal{L}\left\{\int_0^t e^{-t}\,dt\right\} = \frac{1}{s}F(s) = \frac{1}{s(s+1)}$$

6.1.1 The inverse transform

The inverse Laplace transform is the transformation of a Laplace transform into a function of time. If $\mathcal{L}\{f(t)\} = F(s)$ then $f(t)$ is the *inverse Laplace transform* of $F(s)$, the inverse being written as:

$$f(t) = \mathcal{L}^{-1}\{F(s)\} \qquad [13]$$

The inverse can generally be obtained by using standard transforms, e.g. those in Table 6.1. The basic properties of the inverse, see the following notes, can be used with the standard transforms to obtain a wider range of transforms than just those in the table. Often $F(s)$ is the ratio of two polynomials and cannot be readily identified with a standard transform. However, the use of partial fractions (see Section 4.2.3) can often convert such an expression into simple fraction terms which can then be identified with standard transforms. This is illustrated in the examples given in the next section.

Example

Determine the inverse Laplace transform of $1/s^2$.

Table 6.1 indicates that the function which has the Laplace transform of $1/s^2$ is t. Thus the inverse is t.

Basic properties of the inverse transform

The following are basic properties which aid in the obtaining of inverse transforms.

- ***Additive property***
  If we have a Laplace transform as the sum of two separate terms then we can take the inverse of each separately and the sum of the two inverse transforms is the inverse of the sum:

$$\mathcal{L}^{-1}\{F(s) + G(s)\} = \mathcal{L}^{-1}\{F(s)\} + \mathcal{L}^{-1}\{G(s)\} \qquad [14]$$

  Also:

$$\mathcal{L}^{-1}\{aF(s)\} = a\mathcal{L}^{-1}\{F(s)\} \qquad [15]$$

  where a is a constant.

- ***First shift theorem***
  The *first shift theorem* (see Section 6.1) can be written in inverse form as:

$$\mathcal{L}^{-1}\{F(s-a)\} = e^{at}f(t) \qquad [16]$$

  where $f(t)$ is the inverse transform of $F(s)$.

- ***Second shift theorem***
  The *second shift theorem* (see Section 6.1) can be written in inverse form as:

$$\mathcal{L}^{-1}\{e^{-sT}F(s)\} = f(t-T)u(t-T) \qquad [17]$$

Key points

$f(t) = \mathcal{L}^{-1}\{F(s)\}$

A Laplace transform which is the sum of two separate terms has an inverse of the sum of the inverse transforms of each term considered separately.

A Laplace transform which is a constant multiplied by a function has an inverse of the constant multiplied by the inverse of the function.

First shift theorem: $\mathcal{L}^{-1}\{F(s-a)\} = e^{at}f(t)$, where $f(t)$ is the inverse transform of $F(s)$.

Second shift theorem: if the inverse transform numerator contains an e^{-sT} term, we remove this term from the expression, determine the inverse transform of what remains and then substitute $(t-T)$ for t in the result.

Thus if the inverse transform numerator contains an e^{-sT} term, then we remove this term from the expression, determine the inverse transform of what remains and then substitute $(t - T)$ for t in the result.

Example

Determine the inverse Laplace transform of $\frac{7s}{s^2+9}$.

Table 6.1 shows the Laplace transform of cos ωt as being $s/(s^2 + \omega^2)$. Thus:

$$\mathcal{L}^{-1}\left\{\frac{s}{s^2+\omega^2}\right\} = \cos\omega t$$

Thus, using equation [15]:

$$\mathcal{L}^{-1}\left\{\frac{7s}{s^2+9}\right\} = 7\mathcal{L}^{-1}\left\{\frac{s}{s^2+9}\right\} = 7\cos 3t$$

Example

Determine the inverse Laplace transform of $\frac{3s-1}{s(s-1)}$.

We can write the fraction in a simpler form by the use of partial fractions. Thus:

$$\frac{3s-1}{s(s-1)} = \frac{A}{s} + \frac{B}{s-1}$$

and so we must have $3s - 1 = A(s - 1) + Bs$. Equating coefficients of s gives $3 = A + B$ and equating numerical terms gives $-1 = -A$. Hence:

$$\frac{3s-1}{s(s-1)} = \frac{1}{s} + \frac{2}{s-1}$$

The inverse transform of $1/s$ is 1 and of $1/(s - 1)$ is e^t. Thus:

$$\mathcal{L}^{-1}\left\{\frac{3s-1}{s(s-1)}\right\} = 1 + 2\,e^t$$

Example

Determine the inverse Laplace transform of $\frac{6}{s^2-6s+13}$.

This fraction can be rearranged as:

$$\frac{6}{s^2-6s+13}=3\frac{2}{(s-3)^2+2^2}$$

The fraction term is now in the form $\omega/(s^2+\omega^2)$, i.e. the transform of sin ωt when s has been replaced by $s-3$. This corresponds to a multiplication by e^{3t}. Thus, using equation [16]:

$$\mathcal{L}^{-1}\left\{\frac{5}{s^2-6s+13}\right\}=3\ e^{3t}\sin 2t$$

Example

Determine the inverse Laplace transform of $6^{e-3t}/(s+2)$.

Using equation [17], extracting e^{-3s} from the expression gives $6/(s+2)$. This has the inverse Laplace transform of $6\ e^{-2t}$. Thus the required inverse is $5(t-3)\ e^{-2(t-3)}u(t-3)$.

Initial and final values

The *initial value* of a function of time is its value at zero time, the *final value* being the value at infinite time. Often there is a need to determine the initial value and final values of systems, e.g. for an electrical circuit when there is, say, a step input. The final value in such a situation is often referred to as the *steady-state value*. The initial and final value theorems enable the initial and final values to be determined from a Laplace transform without the need to find the inverse transform.

- ***The initial value theorem***

  The Laplace transform of $f(t)$ is given by equation [1] as:

  $$\mathcal{L}\{f(t)\}=\int_0^{\infty}e^{-st}f(t)\,dt$$

  and so:

  $$\mathcal{L}\left\{\frac{d}{dt}f(t)\right\}=\int_0^{\infty}e^{-st}\frac{d}{dt}f(t)\,dt \qquad [18]$$

  Integration by parts then gives:

  $$\mathcal{L}\left\{\frac{d}{dt}f(t)\right\}=\left[e^{-st}f(t)\right]_0^{\infty}-\int_0^{\infty}(-s\,e^{-st})f(t)\,dt=-f(0)+sF(s \qquad [19]$$

  As s tends to infinity then e^{-st} tends to 0. Thus we must have, as a result of equation [18], $\mathcal{L}\{df(t)/dt\}$ tending to 0 as s tends to infinity. Hence equation [19] gives:

$$\lim_{s\to\infty}[-f(0)+sF(s)]=0$$

$$\lim_{s\to\infty}sF(s)=f(0)$$

Key point

Initial value theorem:

$\lim_{t\to 0} f(t) = \lim_{s\to\infty} sF(s)$

But $f(0)$ is the initial value of the function at $t = 0$. Thus, provided a limit exists:

$$\lim_{t\to 0}f(t)=\lim_{s\to\infty}sF(s) \qquad [20]$$

This is known as the *initial value theorem*.

Example

Determine the initial value of the function $f(t)$ giving the Laplace transform $4/(s + 2)$.

Applying equation [20]:

$$\lim_{t\to 0}f(t)=\lim_{s\to\infty}\left[\frac{4s}{s+2}\right]=\lim_{s\to\infty}\left[\frac{4}{1+2/s}\right]=4$$

- ***The final value theorem***

As with the initial value theorem, for a function $f(t)$ having a Laplace transform $F(s)$ we can write (equations [18] and [19]):

$$\mathcal{L}\left\{\frac{\mathrm{d}}{\mathrm{d}t}f(t)\right\}=\int_0^\infty \mathrm{e}^{-st}\frac{\mathrm{d}}{\mathrm{d}t}f(t)\,\mathrm{d}t=-f(0)+sF(s) \qquad [21]$$

As s tends to zero then e^{-st} tends to 1 and so:

$$\lim_{s\to 0}\left[\int_0^\infty \mathrm{e}^{-st}\frac{\mathrm{d}}{\mathrm{d}t}f(t)\,\mathrm{d}t\right]=\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}f(t)\,\mathrm{d}t$$

We can write this integral as:

$$\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}f(t)\,\mathrm{d}t=\lim_{t\to\infty}\int_0^t \frac{\mathrm{d}}{\mathrm{d}t}f(t)\,\mathrm{d}t=\lim_{t\to\infty}\left[f(t)-f(0)\right]$$

Hence, with equation [21] we obtain:

$$\lim_{s\to 0}\left[-f(0)+sF(s)\right]=\lim_{t\to\infty}\left[f(t)-f(0)\right]$$

Key point

Final value theorem:

$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s)$

and so, provided a limit exists:

$$\lim_{t\to\infty}f(t)=\lim_{s\to 0}sF(s) \qquad [22]$$

This is termed the *final value theorem*.

Example

Determine the final value of the function which has the Laplace transform:

$$F(s) = \frac{2s+1}{(s+1)(s+3)}$$

Using equation [22]:

$$\lim_{t\to\infty} f(t) = \lim_{s\to\infty} sF(s) = \lim_{s\to\infty}\left[\frac{s(2s+1)}{(s+1)(s+3)}\right] = 0$$

Problems 6.1

1 Determine, working from first principles and the definition of the transform, the Laplace transforms of:

(a) $f(t) = t^2$, (b) $f(t) = t^3$, (c) $f(t) = \sinh at$.
(Hint: $\sinh at = \frac{1}{2}(e^{at} - e^{-at})$)

2 Determine, by the use of the transforms given in Table 6.1 and the properties of Laplace transforms, the Laplace transforms of the following functions:
(a) 4, (b) $3t - 1$, (c) e^{3t}, (d) $2t + 3\,e^t$, (e) $t^2 + 4\,e^{-2t}$,
(f) $t^2 + 2t + 1$, (g) $2 \sin 3t$, (h) $5 \sinh 3t$, (i) $\sin 3t \cos 3t$,
(j) $t\,e^{-3t}$, (k) $4 - 2\sin 3t + e^{2t}$, (l) $t^3\,e^{-2t}$,
(m) $(1 + e^t)(1 - e^{-t})$, (n) $e^{3t}\cos t$, (o) $(1 + t)^2\,e^{-t}$,
(p) $e^{-t}\sin^2 t$, (q) $t\cosh 3t$, (r) $t^2\cosh 3t$, (s) $t^3\,e^{-3t}$

3 Use the additive property to determine the Laplace transforms of the following functions:

(a) $t^2 + 3t + 2$, (b) $2 + 4\sin 3t$, (c) $e^{4t} + \cosh 2t$,
(d) $2 + 5\,e^{3t}$, (e) $\cos 2t + \cos 3t$, (f) $t^3 + 4\,e^{-t}$

4 Use the first shift theorem to determine the Laplace transforms of the following functions:

(a) $e^{-3t}\sin 2t$, (b) $e^{4t}t^2$, (c) $e^{2t}\cos t$

5 Use the second shift theorem to determine the Laplace transform of the following functions:

(a) a unit step function which starts at $t = 5$ s,
(b) a unit impulse which occurs at $t = 4$ s,
(c) the function described by $3(t - 10)u(t - 10)$

6 Determine the Laplace transform of the periodic function shown in Figure 6.8.

7 Determine the Laplace transform for the periodic signal shown in Figure 6.9.

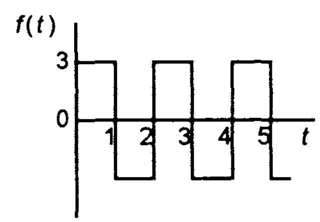


Figure 6.8 *Problem 6*

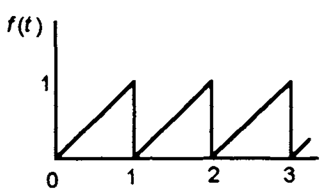


Figure 6.9 *Problem 7*

8 Determine the Laplace transform for the following periodic signals:

(a) $f(t) = 1$ for $0 \le t < 1$ and 0 for $1 \le t < 2$, $f(t + 2) = f(t)$,
(b) $f(t) = t$ for $0 \le t < 1$ and 0 for $1 \le t < 2$, $f(t + 2) = f(t)$,
(c) $f(t) = t$ for $0 \le t < 1$ and $2 - t$ for $1 \le t < 2$, $f(t + 2) = f(t)$

9 Determine the inverse Laplace transforms of:

(a) $\frac{1}{s-2}$, (b) $\frac{5}{s}$, (c) $\frac{s}{s^2+16}$, (d) $\frac{3}{s^2-9}$,

(e) $\frac{5}{(s-2)^2+25}$, (f) $\frac{1}{(s+3)^4}$, (g) $\frac{e^{-2s}}{s^2}$, (h) $\frac{e^{-3s}}{(s+2)^2}$

10 Determine, by the use of partial fractions, the inverse Laplace transforms of the following:

(a) $\frac{3s+1}{s^2-s-6}$, (b) $\frac{3s+3}{s^2+s-2}$, (c) $\frac{s-4}{(s+1)(s^2+4)}$

(d) $\frac{s+4}{s^2+4s+4}$

11 Determine the initial values of the functions giving the following Laplace transforms:

(a) $\frac{1}{s^2+1}$, (b) $\frac{s}{s^2+1}$

12 Determine the final values of the functions having the following Laplace transforms:

(a) $\frac{2}{s}$, (b) $\frac{1}{s+5}$

6.2 Solving differential equations

Laplace transforms offer a method of solving differential equations. The procedure adopted is:

1 Replace each term in the differential equation by its Laplace transform, inserting the given initial conditions.

2 Algebraically rearrange the equation to give the transform of the solution.

3 Invert the resulting Laplace transform to obtain the answer as a function of time.

Example

Given that $x = 0$ at $t = 0$, solve the first-order differential equation $3(dx/dt) + 2x = 4$.

Taking the Laplace transform gives:

$$3[sX(s) - x(0)] + 2X(s) = \frac{4}{s}$$

Substituting the initial condition gives:

$$3sX(s) + 2X(s) = \frac{4}{s}$$

Hence:

$$X(s) = \frac{4}{s(3s+2)}$$

Simplifying by the use of partial fractions:

$$\frac{4}{s(3s+2)} = \frac{A}{s} + \frac{B}{3s+2}$$

Hence $A(3s + 2) + Bs = 4$ and so $A = 2$ and $B = -2/3$. Thus:

$$X(s) = \frac{2}{s} - \frac{2}{3(3s+2)} = \frac{2}{s} - 2\frac{\frac{2}{3}}{\frac{2}{3}\left(s+\frac{2}{3}\right)}$$

and so $x(t) = 2 - 2\,e^{-2t/3}$.

Example

Given that $x = 0$ and $dx/dt = 1$ at $t = 0$, solve the second-order differential equation:

$$\frac{d^2x}{dt^2} - 5\frac{dx}{dt} + 6x = 2\,e^{-t}$$

Taking the Laplace transform gives:

$$s^2X(s) - sx(0) - \frac{d}{dt}x(0) - 5[sX(s) - x(0)] + 6X(s) = \frac{2}{s+1}$$

Substituting the initial conditions:

$$s^2X(s) - 1 - 5sX(s) + 6X(s) = \frac{2}{s+1}$$

$$X(s) = \frac{\frac{2}{s+1}+1}{s^2-5s+6} = \frac{2}{(s+1)(s-2)(s-3)} + \frac{1}{(s-2)(s-3)}$$

We can simplify the above expression by the use of partial fractions. Thus:

$$\frac{2}{(s+1)(s-2)(s-3)} = \frac{A}{s+1} + \frac{B}{s-2} + \frac{C}{s-3}$$

Hence $A(s-2)(s-3) + B(s+1)(s-3) + C(s+1)(s-2) = 2$ and so $A = 1/6$, $B = -2/3$ and $C = ½$.

$$\frac{1}{(s-2)(s-3)} = \frac{D}{s-2} + \frac{E}{s-3}$$

Hence $D(s-3) + E(s-2) = 1$ and so $D = -1$ and $E = 1$. Thus:

$$X(s) = \frac{\frac{1}{6}}{s+1} + \frac{-\frac{2}{3}}{s-2} + \frac{\frac{1}{2}}{s-3} + \frac{-1}{s-2} + \frac{1}{s-3}$$

$$= \frac{\frac{1}{6}}{s+1} - \frac{\frac{5}{3}}{s-2} + \frac{\frac{3}{2}}{s-3}$$

The inverse transform is $x(t) = \frac{1}{6}\,e^{-t} - \frac{5}{3}\,e^{2t} + \frac{3}{2}\,e^{3t}$.

Example

Solve the following second-order differential equation:

$$\frac{d^2x}{dt^2} + 64x = 0$$

given the conditions (a) $dx/dt = 0$ and $x = 2$ when $t = 0$, (b) $dx/dt = 2$ and $x = 0$ when $t = 0$.

Taking the Laplace transform gives:

$$s^2X(s) - sx(0) - \frac{d}{dt}x(0) + 64X(s) = 0$$

(a) We have $dx/dt = 0$ and $x = 2$ when $t = 0$ and so:

$$s^2X(s) - 2s - 0 + 64X(s) = 0$$

$$(s^2 + 64)X(s) = 2s$$

$$X(s) = \frac{2s}{s^2+64} = 2\left(\frac{s}{s^2+8^2}\right)$$

As Table 6.1 indicates, the bracketed term has the inverse of a cosine. Thus the solution is $x = 2\cos 8t$.

(b) We have $dx/dt = 2$ and $x = 0$ when $t = 0$, and so:

$$s^2X(s) - 0 - 2 + 64X(s) = 0$$

$$(s^2 + 64)X(s) = 2$$

$$X(s) = \frac{2}{s^2+64}$$

To put this in the form $\omega/(s^2 + \omega^2)$ we multiply by 4/4:

$$X(s) = \frac{1}{4}\frac{8}{s^2+8^2}$$

The solution is thus $x = \frac{1}{4}\sin 8t$.

Problems 6.2

1 Solve the following differential equations:

(a) $2\frac{dx}{dt} + x = 4\,e^{2t}$,
with $x = 0$ when $t = 0$,

(b) $\frac{dx}{dt} + 5x = 2$,
with $x = 0$ when $t = 0$,

(c) $\frac{d^2x}{dt^2} + 4x = 1$,
with $x = 0$ and $dx/dt = 0$ when $t = 0$,

(d) $\frac{d^2x}{dt^2} + 2\frac{dx}{dt} + 2x = e^{-t}$,
with $x = 0$ and $dx/dt = 0$ when $t = 0$,

(e) $\frac{d^2x}{dt^2} - 2\frac{dx}{dt} + x = \sin t$, with $x = 1$
with $x = 1$ and $dx/dt = 0$ when $t = 0$,

(f) $\frac{dx}{dt} - 2x = 3$,
with $x = 0$ when $t = 0$,

(g) $\frac{dx}{dt} + 4x = \cos t$,
with $x = 0$ when $t = 0$,

(h) $\frac{d^2x}{dt^2} + x = 3$,
with $x = 0$ and $dx/dt = 1$ when $t = 0$,

(i) $\frac{d^2x}{dt^2} + \frac{dx}{dt} - 6x = 5\,e^{3t}$
with $x = 0$ and $dx/dt = 0$ when $t = 0$,

(j) $\frac{d^2x}{dt^2} - \frac{dx}{dt} - 6x = \cos 2t$
with $x = 0$ and $dx/dt = 0$ when $t = 0$,

(k) $\frac{d^2x}{dt^2} - 5\frac{dy}{dt} - 6x = t + e^{3t}$
with $x = 2$ and $dx/dt = 1$ when $t = 0$.

6.3 Transfer function

In general, when we consider inputs and outputs of systems as functions of time then the relationship between the output and input is given by a differential equation. If we have a system composed of two elements in series with each having its input–output relationships described by a differential equation, it is not easy to see how the output of the system as a whole is related to its input. We can overcome this problem by transforming the differential equations into a more convenient form by using the Laplace transform. This form is a much more convenient way of describing the relationship than a differential equation since it can be easily manipulated by the basic rules of algebra.

For a simple system we might use the term gain to relate the input and output of a system with gain G = output/input. This tells us how much bigger the output is than the input. When we are working with inputs and outputs described as functions of s we define the *transfer function* $G(s)$ as [output $Y(s)$]/[input $X(s)$] when all initial conditions before we apply the input are zero, i.e.

Key point

The *transfer function* $G(s)$ of a system is defined as [output $Y(s)$]/[input $X(s)$] when all initial conditions before we apply the input are zero.

$$G(s) = \frac{Y(s)}{X(s)} \qquad [23]$$

Figure 6.10 *Transfer function as the factor that multiplies the input to give the output*

A transfer function can be represented as a block diagram (Figure 6.10) with $X(s)$ the input, $Y(s)$ the output and the transfer function $G(s)$ as the operator in the box that converts the input to the output. The block represents a multiplication for the input. Thus, by using the Laplace transform of inputs and outputs, we can use the transfer function as a simple multiplication factor, like the gain.

Example

Determine the transfer function for an electrical system for which we have the relationship:

$$\frac{V_C(s)}{V(s)} = \frac{1}{RCs+1}$$

The transfer function $G(s)$ is thus:

$$G(s) = \frac{V_C(s)}{V(s)} = \frac{1}{RCs+1}$$

To get the output $V_C(s)$ we multiply the input $V(s)$ by $1/(RCs + 1)$.

Example

Determine the transfer function for the mechanical system having mass, stiffness and damping, and input F and output x and described by the differential equation:

$$F = m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx$$

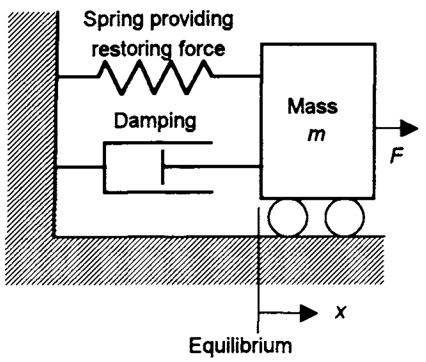


Figure 6.11 *Mass, spring, damper system*

Figure 6.11 shows the type of system that would give such a differential equation.

If we now write $\mathcal{L}(x) = X(s)$ and $\mathcal{L}(F) = F(s)$, and with initial conditions zero:

$$\mathcal{L}\left\{m\frac{d^2x}{dt^2}\right\} = m\left[s^2X(s) - sx(0) - \frac{d}{dt}x(0)\right]$$

$$\mathcal{L}\left\{c\frac{dx}{dt}\right\} = c\left[sX(s) - x(0)\right]$$

$$\mathcal{L}\{k(x)\} = kX(s)$$

$$\mathcal{L}\{F\} = F(s)$$

When $t = 0$ we have $x = 0$ and $dx/dt = 0$ and so:

$$\mathcal{L}\left\{m\frac{d^2x}{dt^2}\right\} = m\left[s^2X(s)\right]$$

$$\mathcal{L}\left\{c\frac{dx}{dt}\right\} = c\left[sX(s)\right]$$

Thus, we have for the differential equation in the s-domain:

$$F(s) = ms^2X(s) + csX(s) + kX(s)$$

$$= (ms^2 + cs + k)X(s)$$

Hence the transfer function $G(s)$ of the system is:

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{ms^2 + cs + k}$$

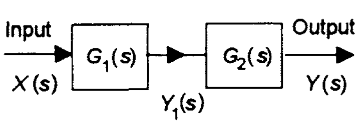


Figure 6.12 *Systems in series*

Key point

The overall transfer function for a system composed of elements in series is the product of the transfer functions of the individual series elements.

Systems in series

Consider a system of two subsystems in series (Figure 6.12). The first subsystem has an input of $X(s)$ and an output of $Y_1(s)$; thus, $G_1(s) = Y_1(s)/X(s)$. The second subsystem has an input of $Y_1(s)$ and an output of $Y(s)$; thus, $G_2(s) = Y(s)/Y_1(s)$. We thus have:

$$Y(s) = G_2(s)Y_1(s) = G_2(s)G_1(s)X(s)$$

The overall transfer function $G(s)$ of the system is $Y(s)/X(s)$ and so:

$$G_{overall}(s) = G_1(s)G_2(s) \qquad [24]$$

Example

Determine the overall transfer function for a system which consists of two elements in series, one having a transfer function of $1/(s + 1)$ and the other $1/(s + 2)$.

The overall transfer function is thus:

$$G_{\text{overall}}(s) = \frac{1}{s+1} \times \frac{1}{s+2} = \frac{1}{(s+1)(s+2)}$$

Key point

A simple feedback control system to, say, control the temperature of a room will have a negative feedback loop. This feeds back a measure of the output of the system which is then subtracted from the input. The input is the required temperature and the output the actual temperature. The difference between these signals, i.e. the error, is used to actuate some heating system which will continue as long as there is an error.

Systems with negative feedback

For a control system with a negative feedback loop we can have the situation shown in Figure 6.13 where the output is fed back via a system with a transfer function $H(s)$. This fed back signal subtracts from the input to the system $G(s)$ to give the error signal. The feedback system has an input of $Y(s)$ and thus an output of $H(s)Y(s)$. Thus the feedback signal is $H(s)Y(s)$. The error is the difference between the input signal $X(s)$ and the feedback signal:

$$\text{error}(s) = X(s) - H(s)Y(s)$$

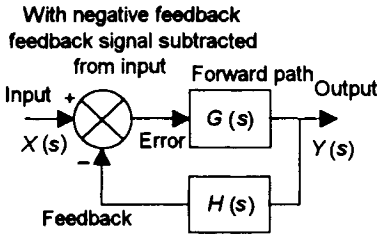


Figure 6.13 *System with negative feedback*

This error signal is the input to the $G(s)$ system and gives an output of $Y(s)$. Thus:

$$G(s) = \frac{Y(s)}{X(s) - H(s)Y(s)}$$

and so:

$$[1 + G(s)H(s)]Y(s) = G(s)X(s)$$

which can be rearranged to give:

$$\text{overall transfer function} = \frac{Y(s)}{X(s)} = \frac{G(s)}{1 + G(s)H(s)} \qquad [25]$$

Key point

For a system with a negative feedback, the overall transfer function is the forward path transfer function divided by one plus the product of the forward path and feedback path transfer functions.

Example

Determine the overall transfer function for a control system (Figure 6.14) which has a negative feedback loop with a transfer function 4 and a forward path transfer function of $2/(s + 2)$.

The overall transfer function of the system is:

$$G_{\text{overall}}(s) = \frac{\frac{2}{s+2}}{1 + 4 \times \frac{2}{s+2}} = \frac{2}{s+10}$$

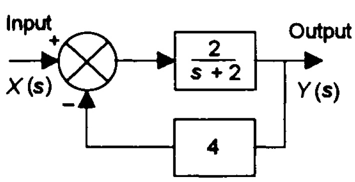


Figure 6.14 *Example*

6.3.1 Determining outputs of systems

The procedure we can use to determine how the output of a system will change with time when there is some input to the system is:

1. ***Determine the output as an s function***
   In terms of the transfer function $G(s)$ we have:

   $$\text{Output }(s) = G(s) \times \text{Input }(s) \qquad [26]$$

   We can thus obtain the output of a system as an s function by multiplying its transfer function by the input s function.

2. ***Determine the time function corresponding to the output s function***
   To obtain the output as a function of time we need to find the time function that will give the particular output s function that we have obtained. Tables of s functions and their corresponding time functions can be used (Table 6.1). Often, however, the s function output has to be rearranged to put it into a form given in the table.

> **Example**
>
> A system has a transfer function of $1/(s + 2)$. What will be its output as a function of time when it is subject to a step input of 1 V?
>
> The step input has a Laplace transform of $(1/s)$. Thus:
>
> Output $(s) = G(s) \times$ Input (s)
>
> $$= \frac{1}{s+2} \times \frac{1}{s} = \frac{1}{s(s+2)}$$
>
> The nearest form we have in Table 6.1 to the output is item 6 as $\frac{1}{2} \times 2/[s(s + 2)]$. Thus the output, as a function of time, is $\frac{1}{2}(1 - e^{-5t})$ V.

First-order systems

A first-order system has a differential equation of the form:

$$\tau \frac{dy}{dt} + y = kx$$

As a function of s this can be written as:

$$\tau Y(s) + Y(s) = kX(s)$$

and so a transfer function of the form:

$$G(s) = \frac{Y(s)}{X(s)} = \frac{k}{\tau s + 1} \tag{27}$$

where k is the *gain* of the system when there are steady-state conditions and τ is the *time constant* of the system.

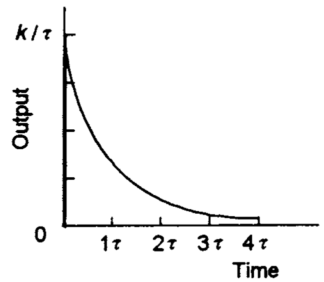


Figure 6.15 *Output with a unit impulse input to a first-order system*

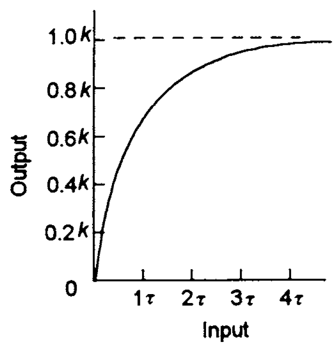


Figure 6.16 *Behaviour of a first-order system when subject to a unit step input*

- ***Unit impulse input***

  When a first-order system is subject to a unit impulse input then $X(s) = 1$ and the output transform $Y(s)$ is:

  $$Y(s) = G(s)X(s) = \frac{k}{\tau s + 1} \times 1 = k\frac{(1/\tau)}{s + 1/\tau}$$

  Hence, since we have the transform in the form $1/(s + a)$, using item 6 in Table 6.1 gives:

  $$x = k(1/\tau)\, e^{-t/\tau} \tag{28}$$

  Figure 6.15 shows how the output x varies with time.

- ***Unit step input***

  When a first-order system is subject to a unit step input then $X(s) = 1/s$ and the output transform $Y(s)$ is:

  $$X(s) = G(s)Y(s) = \frac{k}{s(\tau s + 1)} = k\frac{(1/\tau)}{s(s + 1/\tau)}$$

  Hence, since we have the transform in the form $a/s(s + a)$, using item 6 in Table 6.1 gives:

  $$x = k(1 - e^{-t/\tau}) \tag{29}$$

  Figure 6.16 shows how the output x varies with time.

Example

A circuit has a resistance R in series with a capacitance C. The differential equation relating the input v and output v_C, i.e. the voltage across the capacitor, is:

$$v = RC\frac{dv_C}{dt} + v_C$$

Determine the output of the system when there is a 2 V impulse input.

As a function of s the differential equation becomes:

$$V(s) = RCsV_C(s) + V_C(s)$$

Hence the transfer function is

$$G(s) = \frac{V_C(s)}{V(s)} = \frac{1}{RCs+1}$$

The output when there is 2 V impulse input is:

$$V_C(s) = G(s)V(s) = \frac{1}{RCs+1} \times 2 = \frac{2/RC}{s+1/RC}$$

Hence, since we have the transform in the form $1/(s + a)$, using item 6 in Table 6.1 gives:

$v = (2/RC)\, e^{-t/RC}$

Example

A thermocouple which has a transfer function linking its voltage output V and temperature input of:

$$G(s) = \frac{30 \times 10^{-6}}{10s+1}\ \text{V/°C}$$

Determine the response of the system when it is suddenly immersed in a water bath at 100°C.

The output as an s function is:

$V(s) = G(s) \times \text{input}\ (s)$

The sudden immersion of the thermometer gives a step input of size 100°C and so the input as an s function is $100/s$. Thus:

$$V(s) = \frac{30 \times 10^{-6}}{10s+1} \times \frac{100}{s} = \frac{30 \times 10^{-4}}{10s(s+0.1)}$$

$$= 30 \times 10^{-4} \frac{0.1}{s(s+0.1)}$$

The fraction element is of the form $a/s(s + a)$, item 6 in Table 6.1, and so the output as a function of time is:

$V = 30 \times 10^{-4}\,(1 - e^{-0.1t})$ V

Second-order systems

The differential equation for a second-order system can be written as:

$$\frac{d^2y}{dt^2} + 2\zeta\omega_n\frac{dy}{dt} + \omega_n^2 y = kx$$

where x is the input and y the output (see Chapter 5, equation [36]). Since the steady-state output occurs when $\omega_n^2 y = kx$, a more usual way of writing the standard form of the equation, so that the steady-state value occurs when $y = kx$, is as:

$$\frac{d^2y}{dt^2} + 2\zeta\omega_n\frac{dy}{dt} + \omega_n^2 y = k\omega_n^2 x \quad [30]$$

where ω_n is the natural angular frequency with which the system oscillates and ζ is the damping ratio. Hence we have:

$$s^2Y(s) + 2\zeta\omega_n sY(s) + \omega_n^2 Y(s) = k\omega_n^2 X(s)$$

and so a transfer function of:

$$G(s) = \frac{Y(s)}{X(s)} = \frac{k\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad [31]$$

When a second-order system is subject to a unit step input, i.e. $X(s) = 1/s$, then the output transform is:

$$Y(s) = G(s)X(s) = \frac{k\omega_n^2}{s(s^2 + 2\zeta\omega_n s + \omega_n)}$$

Key point

In general, we can write a transfer function as:

$$G(s) = \frac{K(s-z_1)(s-z_2)\dots(s-z_m)}{(s-p_1)(s-p_2)\dots(s-p_n)}$$

with the values of s that make $G(s)$ zero being termed zeros and so correspond to $s = z_1, z_2, \dots z_m$. The values of s that make $G(s)$ infinite are known as poles and so correspond to $s = p1, p2, \dots p_n$. As will be apparent from the discussion on this page and the next, poles and zeros can be real or complex. We can thus plot them on an Argand diagram with the vertical axis being the imaginary element and the horizontal axis the real part. Such a plot is said to have the poles and zeros plotted on an s-plane diagram.

There are three different forms of answer to this equation for the way the output varies with time; these depending on the value of the damping constant and whether it gives an overdamped, critically damped or underdamped system. We can determine the condition for these three forms of output by putting the equation in the form:

$$Y(s) = \frac{k\omega_n^2}{s(s+p_1)(s+p_2)} \quad [32]$$

where p_1 and p_2 are the roots of the quadratic term:

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = 0 \quad [33]$$

Hence, if we use the equation to determine the roots of a quadratic equation, we obtain:

$$p = \frac{-2\zeta\omega_n \pm \sqrt{4\zeta^2\omega_n^2 - 4\omega_n^2}}{2}$$

and so the two roots are given by:

$$p_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1} \text{ and } p_2 = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1} \quad [34]$$

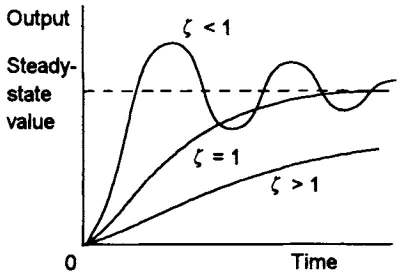


Figure 6.17 *Behaviour of a second-order system when subject to a unit step input signal*

The important issue in determining the form of the roots is the value of the square root term and this is determined by the value of the damping factor (Figure 6.17).

- ***Damping factor*** $\zeta > 1$

With the damping factor ζ greater than 1 the square root term is real and will factorise. To find the inverse transform we can either use partial fractions to break the expression down into a number of simple fractions or use item 10 in Table 6.1. The output is thus:

$$y = \frac{k\omega_n^2}{p_1 p_2}\left[1 - \frac{p_2}{p_2 - p_1}\mathrm{e}^{-p_1 t} + \frac{p_1}{p_2 - p_1}\mathrm{e}^{-p_2 t}\right] \qquad [35]$$

This describes an output which does not oscillate but dies away with time and thus the system is *overdamped*. As the time t tends to infinity then the exponential terms tend to zero and the output becomes the steady value of $k\omega_n^2/(p_1p_2)$. Since $p_1p_2 = \omega_n^2$, the steady value is k.

- ***Damping factor*** $\zeta = 1$

With $\zeta = 1$ the square root term is zero and so $p_1 = p_2 = -\omega_n$; both roots are real and both the same. The output equation then becomes:

$$Y(s) = \frac{k\omega_n^2}{s(s+\omega_n)^2}$$

This equation can be expanded by means of partial fractions to give:

$$Y(s) = k\left[\frac{1}{s} - \frac{1}{s+\omega_n} - \frac{\omega_n}{(s+\omega_n)^2}\right]$$

Hence:

$$y = k[1 - \mathrm{e}^{-\omega_n t} - \omega_n t\,\mathrm{e}^{-\omega_n t}] \qquad [36]$$

This is the critically damped condition and describes an output which does not oscillate but dies away with time. As the time t tends to infinity then the exponential terms tend to zero and the output tends to the steady state value of k.

- ***Damping factor*** $\zeta < 1$

With $\zeta < 1$ the square root term does not have a real value. Using item 19 in Table 6.1 then gives:

$$y = k\left[1 - \frac{\mathrm{e}^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\omega_n\sqrt{(1-\zeta^2)}\,t + \phi\right)\right] \qquad [37]$$

where $\cos\phi = \zeta$. This is an under-damped oscillation. The angular frequency ω of the damped oscillation is:

$$\omega = \omega_n \sqrt{1-\zeta^2} \qquad [38]$$

Only when the damping is very small does the angular frequency of the oscillation become nearly the natural angular frequency ω_n. As the time t tends to infinity then the exponential term tends to zero and so the output tends to the value k.

Example

What will be the state of damping of a system having the following transfer function and subject to a unit step input?

$$G(s) = \frac{1}{s^2 + 8s + 16}$$

The output $Y(s)$ from such a system is given by:

$$Y(s) = G(s)X(s)$$

For a unit step input $X(s) = 1/s$ and so the output is given by:

$$Y(s) = \frac{1}{s(s^2 + 8s + 16)} = \frac{1}{s(s+4)(s+4)}$$

The roots of $s^2 + 8s + 16$ are $p_1 = p_2 = -4$. Both the roots are real and the same, hence we have critical damping.

Example

A system has an output y related to the input x by the differential equation:

$$\frac{d^2y}{dt^2} + 5\frac{dy}{dt} + 6y = x$$

What will be the output from the system when it is subject to a unit step input? Initially both the output and input are zero.

We can write the Laplace transform of the equation as:

$$s^2Y(s) + 5sY(s) + 6Y(s) = X(s)$$

The transfer function is thus:

$$G(s) = \frac{Y(s)}{X(s)} = \frac{1}{s^2 + 5s + 6}$$

For a unit step input the output is given by:

$$Y(s) = \frac{1}{s(s^2+5s+6)} = \frac{1}{s(s+3)(s+2)}$$

Because the quadratic term has two real roots, the system is overdamped. We can directly use one of the standard forms given in Table 6.1 or partial fractions to first simplify the expression before using Table 6.1. Using partial fractions:

$$\frac{1}{s(s+3)(s+2)} = \frac{A}{s} + \frac{B}{s+3} + \frac{C}{s+2}$$

Thus, we have $1 = A(s + 3)(s + 2) + Bs(s + 2) + Cs(s + 3)$. When $s = 0$ then $1 = 6A$ and so $A = 1/6$. When $s = -3$ then $1 = 3B$ and so $B = 1/3$. When $s = -2$ then $1 = -2C$ and so $C = -1/2$. Hence we can write the output in the form:

$$Y(s) = \frac{1}{6s} + \frac{1}{3(s+3)} - \frac{1}{2(s+2)}$$

Hence, using Table 6.1 gives:

$$y = 0.17 + 0.33\ e^{-3t} - 0.5\ e^{-2t}$$

Example

A system has the transfer function:

$$G(s) = \frac{9}{s^2+3.6s+9}$$

Determine its natural frequency, the damping ratio and the frequency of the damped oscillation.

If we compare the transfer function with that given in equation [31], i.e.

$$G(s) = \frac{k\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$$

we are led to conclude that $\omega_n^2 = 9$ and so $\omega_n = 3$ rad/s and $f_n = \omega_n/2\pi = 3/2\pi = 0.48$ Hz. The damping ratio is given by $2\zeta\omega_n = 3.6$ and so $\zeta = 3.6/(2 \times 3) = 0.6$; the system is underdamped. Using equation [38], the angular frequency of the undamped oscillation is given by:

$$\omega = \omega_n\sqrt{1-\zeta^2} = 3\sqrt{1-0.6^2} = 2.4 \text{ rad/s}$$

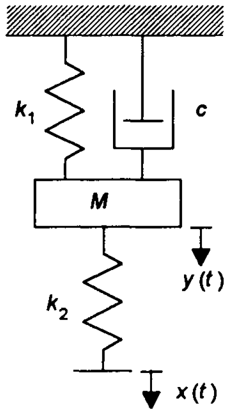


Figure 6.18 *Example*

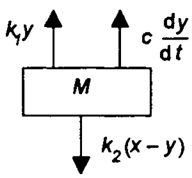


Figure 6.19 *Example*

Example

For the system shown in Figure 6.18, determine its transfer function if M = 50 kg, $k_1 = k_2$ = 400 N/m and c = 180 Ns/m. What will be the damped frequency of its oscillation when subject to a unit step input?

Considering the free-body diagram of the mass (Figure 6.19), and applying Newton's second law, we have:

$$k_2(x-y)-k_1y-c\frac{dy}{dt}=ma=m\frac{d^2y}{dt^2}$$

The Laplace transform of this equation, with zero initial conditions, is:

$$k_2[X(s)-Y(s)]-k_1Y(s)-csY(s)=ms^2Y(s)$$

$$k_2X(s)=ms^2Y(s)+k_2Y(s)+k_1Y(s)+csY(s)$$

$$G(s)=\frac{Y(s)}{X(s)}=\frac{k_2}{ms^2+cs+k_1+k_2}$$

For comparison with the standard form of the transfer function equation, we write the above equation as:

$$G(s)=\frac{(k_2/m)}{s^2+(c/m)s+(k_1+k_2)/m}$$

Hence, with the given data:

$$G(s)=\frac{400/50}{s^2+(180/50)s+(400+400)/50}=\frac{8}{s^2+3.6s+16}$$

Comparing this with the standard form of transfer function for a second-order system [31]:

$$G(s)=\frac{k\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$$

then the natural angular frequency ω_n^2 = 16 and ω_n = 4 rad/s. The damping ratio ζ is given by $2\zeta\omega_n$ = 3.6 and so ζ = 3.6/(2 × 4) = 0.45. The oscillation is underdamped.

Using equation [38], the angular frequency of the undamped oscillation is given by:

$$\omega=\omega_n\sqrt{1-\zeta^2}=4\sqrt{1-0.45^2}=3.57\text{ rad/s}$$

6.3.2 Electrical circuit analysis

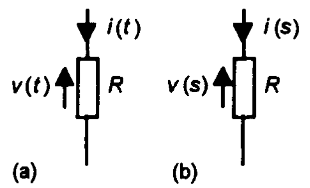


Figure 6.20 *Resistance: (a) time, (b) s-domain*

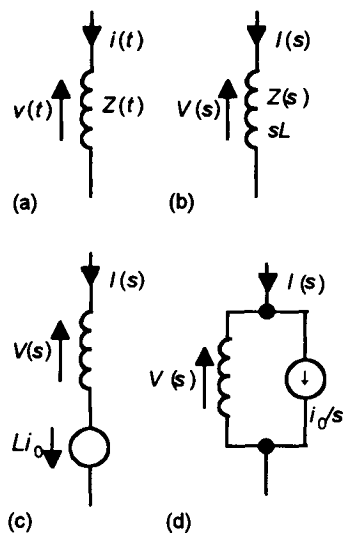


Figure 6.21 *Inductance: (a) time, (b), (c), (d) s-domain*

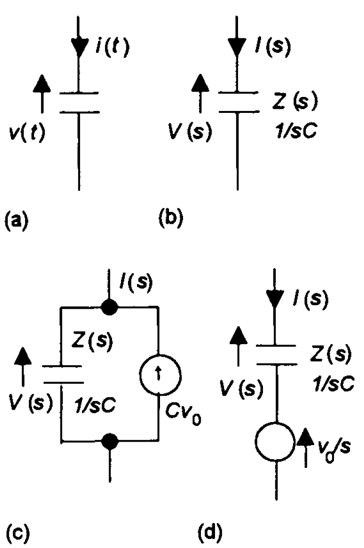


Figure 6.22 *Capacitance: (a) time domain, (b), (c), (d) s-domain*

While we could write differential equations to represent electrical circuits and then solve them by the use of the Laplace transform, a simpler method is to replace time-domain components by their equivalents in the s-domain.

Resistance R in the time domain is defined as $v(t)/i(t)$. Taking the Laplace transform of this equation gives a definition of resistance in the s-domain (Figure 6.20) as:

$$R = \frac{V(s)}{I(s)} \tag{39}$$

Inductance L in the time domain (Figure 6.21(a)) is defined by:

$$v(t) = L\frac{\mathrm{d}i(t)}{\mathrm{d}t}$$

The Laplace transform of this equation is $V(s) = L[sI(s) - i(0)]$. With zero initial current then $V(s) = sLI(s)$. Impedance in the s-domain $Z(s)$ is defined as $V(s)/I(s)$, thus for inductance (Figure 6.21(b)):

$$Z(s) = \frac{V(s)}{I(s)} = sL \tag{40}$$

If the current was not initially zero but $i(0) = i_0$, then $V(s) = sLI(s) - Li_0$. This equation can be considered to describe two series elements (Figure 6.21(c)). The first term then represents the potential difference across the inductance L, being $Z(s)I(s)$, and the second term a voltage generator of $(-Li_0)$. Alternatively we can rearrange equation $V(s) = sLI(s) - Li_0$ in a form to represent two parallel elements (Figure 6.21(d)):

$$I(s) = \frac{V(s) + Li_0}{sL} = \frac{V(s)}{sL} + \frac{i_0}{s} \tag{41}$$

$I(s)$ is the current into the system, $V(s)/sL = V(s)/Z(s)$ can be considered to be the current through the inductance and i_0/s a parallel current source.

Capacitance C in the time domain (Figure 6.22(a)) is defined by:

$$i(t) = C\frac{\mathrm{d}v(t)}{\mathrm{d}t}$$

The Laplace transform of this equation is $I(s) = C[sV(s) - v(0)]$. If we have $v(0) = 0$ then (Figure 6.22(b)):

$$Z(s) = \frac{V(s)}{I(s)} = \frac{1}{sC} \tag{42}$$

If $v(0) = v_0$ then $I(s) = CsV(s) - Cv(0) = CsV(s) - Cv_0$. We can think of this representing $I(s)$ entering a parallel arrangement

Key points

In the *s*-domain:

Resistance = R

Impedance of an inductance = SL

Impedance of a capacitance = $1/sC$

(Figure 6.22(c)) of a capacitor, and giving a current through it is $V(s)/Z(s) = CsV(s)$, and a current source ($-Cv_0$). Alternatively we can rearrange the equation as:

$$V(s) = \frac{1}{sC}I(s) + \frac{v_0}{s} \tag{43}$$

This equation now represents a capacitor in series with a voltage source of v_0/s (Figure 6.22(d)).

Example

Determine the impedance and equivalent series circuit in the *s*-domain of an inductance of 50 mH if there is a current of 0.1 A at time $t = 0$.

The impedance in the *s*-domain is given by equation [40] as $0.050s\ \Omega$. Its equivalent series circuit with the initial condition $i(0) = 0.1$ A is of a voltage source of $-0.050 \times 0.1 = 0.005$ V in series with the impedance of $0.050s\ \Omega$.

Example

Determine the impedance in the *s*-domain of a capacitance of 0.1 μF and its equivalent series circuit when the capacitor has been charged to 5 V at time $t = 0$.

The impedance in the *s*-domain is given by equation [42] as $1/sC = 1/(0.1 \times 10^{-6}s)\ \Omega$, and its equivalent series circuit with the initial condition $v(0) = 5$ V is of a voltage source of $-5/s$ in series with the impedance of $10^7/s\ \Omega$.

Using Kirchhoff's laws

Because of the additive property of the Laplace transform, the transform of a number of time-domain functions is the sum of the transforms of each separate function. Thus with *Kirchhoff's current law*, the algebraic sum of the time-domain currents at a junction is zero and so the sum of the transformed currents is also zero. With *Kirchhoff's voltage law*, the sum of the time-domain voltages around a closed loop is zero and thus the sum of the transformed voltages is also zero. A consequence of this is that:

> *All the techniques developed for use in the analysis of circuits in the time domain can be used in the s-domain.*

Key point

All the techniques developed for use in the analysis of circuits in the time domain can be used in the *s*-domain.

The following examples illustrate this.

Example

Determine the impedance in the s-domain of a 10 Ω resistor in (a) series and (b) parallel with a 1 mH inductor.

(a) For impedances in series $Z(s) = Z_1(s) + Z_2(s) = 10 + 0.001s$ Ω.
(b) For impedances in parallel we have:

$$\frac{1}{Z(s)} = \frac{1}{Z_1(s)} + \frac{1}{Z_2(s)} = \frac{1}{10} + \frac{1}{0.001s} = \frac{0.001s+10}{0.01s}$$

Hence $Z(s) = 0.01s/(0.001s + 10)$ Ω.

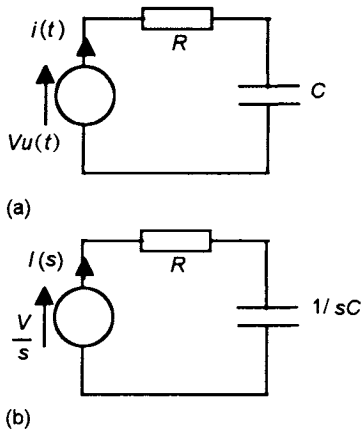


Figure 6.23 *Example*

Example

Determine how the circuit current varies with time for a circuit having a resistance R in series with an initially uncharged capacitance C when the input to the circuit is a step voltage V at time $t = 0$.

Figure 6.23(a) shows the circuit in the time domain and Figure 6.23(b) the equivalent circuit in the s-domain. A unit step at $t = 0$ has the Laplace transform $1/s$ and thus a voltage step of V has a transform of V/s. The impedance of the capacitance is $1/sC$. Thus, applying Kirchhoff's voltage law to the circuit:

$$\frac{V}{s} = RI(s) + \frac{1}{sC}I(s)$$

$$I(s) = \frac{V}{Rs + 1/C} = \frac{V(1/R)}{s + (1/RC)}$$

This is a constant multiplied by $1/(s + a)$, thus:

$$i(t) = \frac{V}{R}\,e^{-t/RC}$$

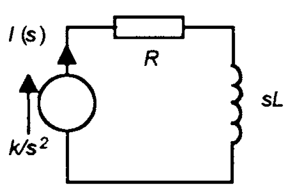


Figure 6.24 *Example*

Example

A ramp voltage of $v = kt$ is applied at time $t = 0$ to a circuit consisting of an inductance L in series with a resistance R. If initially at $t = 0$ there is no current in the circuit, determine how the circuit current varies with time.

The Laplace transform of kt is k/s^2. The inductance has an impedance in the s-domain of sL. Thus the circuit in the s-domain is as shown in Figure 6.24. Applying Kirchhoff's voltage law to the circuit gives:

$$\frac{k}{s^2} = sLI(s) + RI(s)$$

and so:

$$I(s) = \frac{k}{s^2(sL+R)} = \frac{(k/R)(R/L)}{s^2(s+R/L)}$$

This can be simplified by partial fractions, writing a for R/L:

$$\frac{a}{s^2(s+a)} = \frac{A}{s^2} + \frac{B}{s} + \frac{C}{s+a} = \frac{A(s+a)+Bs(s+a)+Cs^2}{s^2(s+a)}$$

Hence $A = 1$, $B = -1/a$ and $C = 1/a$. Thus:

$$I(s) = \frac{k}{R}\left(\frac{1}{s^2} - \frac{1}{(R/L)s} + \frac{1}{(R/L)(s+R/L)}\right)$$

Hence:

$$i(t) = \frac{k}{R}\left(t - \frac{1}{R/L} + \frac{e^{-Rt/L}}{R/L}\right)$$

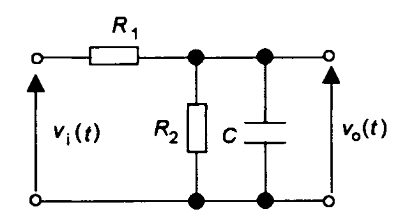


Figure 6.25 *Example*

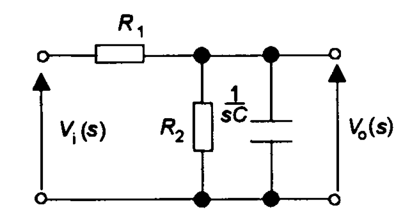


Figure 6.26 *Example*

Example

Determine the transfer function of the circuit shown in Figure 6.25 and the output $v_0(t)$ resulting from a unit step input, given that $R_1 = 10\ \text{k}\Omega$, $R_2 = 22\ \text{k}\Omega$ and $C = 1\ \mu\text{F}$.

The Laplace equivalent circuit is shown in Figure 6.26. The impedance $Zp(s)$ for the parallel arrangement of R_2 and the capacitor is given by:

$$\frac{1}{Z_p(s)} = \frac{1}{R_2} + \frac{1}{1/sC}$$

$$Z_p(s) = \frac{R_2}{1+sCR_2}$$

We have a potential divided circuit and so:

$$V_0(s) = V_i(s)\frac{\left(\dfrac{R_2}{1+sCR_2}\right)}{\left(R_1 + \dfrac{R_2}{1+sCR_2}\right)} = V_i(s)\frac{R_2}{R_1(1+sCR_2)+R_2}$$

The transfer function $G(s)$ of the system is thus:

$$G(s) = \frac{V_0(s)}{V_i(s)} = \frac{R_2}{R_1(1+sCR_2)+R_2}$$

Using the given values:

$$G(s) = \frac{22\times10^3}{220s+32\times10^3}$$

For a unit step input we have $V_i(s) = 1/s$ and so:

$$V_0(s) = \frac{22\times10^3}{220s+32\times10^3}\times\frac{1}{s} = \frac{100}{s(s+145.45)}$$

We can use equation [6] in Table 6.1 or partial fractions to obtain the inverse. Partial fractions give:

$$V_0(s) = \frac{0.687}{s} - \frac{0.687}{s+145.45}$$

and so:

$v_0(t) = 0.687(1 - e^{-145.45t})$

Problems 6.3

1 A system has an input of a voltage of 3 V which is suddenly applied by a switch being closed. What is the input as an s function?

2 A system has an input of a voltage impulse of 2 V. What is the input as an s function?

3 A system has an input of a voltage of a ramp voltage which increases at 5 V per second. What is the input as an s function?

4 A system gives an output of 1/(s + 5) $V(s)$. What is the output as a function of time?

5 A system has a transfer function of $5/(s + 3)$. What will be its output as a function of time when subject to (a) a unit step input of 1 V, (b) a unit impulse input of 1 V?

6 A system has a transfer function of $2/(s + 1)$. What will be its output as a function of time when subject to (a) a step input of 3 V, (b) an impulse input of 3 V?

7 A system has a transfer function of $1/(s + 2)$. What will be its output as a function of time when subject to (a) a step input of 4 V, (b) a ramp input unit impulse of 1 V/s?

8 Use partial fractions to simplify the following expressions:

(a) $\frac{s-6}{(s-1)(s-2)}$, (b) $\frac{s+5}{s^2+3s+2}$, (c) $\frac{2s-1}{(s+1)^2}$

9 A system has a transfer function of:

$$\frac{8(s+3)(s+8)}{(s+2)(s+4)}$$

What will be the output as a time function when it is subject to a unit step input? Hint: use partial fractions.

10 A system has a transfer function of:

$$G(s) = \frac{8(s+1)}{(s+2)^2}$$

What will be the output from the system when it is subject to a unit impulse input? Hint: use partial fractions.

11 What will be the state of damping of systems having the following transfer functions and subject to a unit step input?

$$\text{(a) } \frac{1}{s^2+2s+1}, \text{ (b) } \frac{1}{s^2+7s+12}, \text{ (c) } \frac{1}{s^2+s+1}$$

12 The input x and output y of a system are described by the differential equation:

$$\frac{dy}{dt} + 2y = x$$

Determine how the output will vary with time when there is an input which starts at zero time and then increases at the constant rate of 6 units/s. The initial output is zero.

13 The input x and output y of a system are described by the differential equation:

$$\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + 2y = x$$

If initially the input and output are zero, what will be the output when there is a unit step input?

14 The input x and output y of a system are described by the differential equation:

$$\frac{d^2y}{dt^2} + 4\frac{dy}{dt} + 3y = x$$

If initially the input and output are zero, what will be the output when there is a unit impulse input?

15 A control system has a forward path transfer function of $2/(s + 2)$ and a negative feedback loop with transfer function 4. What will be the response of the system to a unit step input?

16 A system has a transfer function of $100/(s^2 + s + 100)$. What will be its natural frequency ω_n and its damping ratio ζ?

17 A system has a transfer function of $10/(s^2 + 4s + 9)$. Is the system under-damped, critically damped or over-damped?

18 A system has a transfer function of $3/(s^2 + 6s + 9)$. Is the system under-damped, critically damped or over-damped?

19 A system has a forward path transfer function of $10/(s + 3)$ and a negative feedback loop with transfer function 5. What is the time constant of the resulting first-order system?

20 Determine the series and parallel models in the *s*-domain for (a) an inductance of 10 mH when $i(0) = 0.2$ A, (b) a capacitance of 2 μF when $v(0) = 5$ V.

21 Determine the impedance in the *s*-domain of a resistance of 10 Ω in (a) series, (b) parallel with a 2 mH inductance.

22 Determine how the current varies with time when a charged capacitor, with a potential difference of v_0, is allowed to discharge through a resistance R.

23 Determine how the current varies with time when a step voltage $Vu(t)$ is applied to a circuit consisting of a resistance R in series with an inductance L, there being no initial current in the circuit.

24 Determine how the current varies with time when a 1 V impulse is applied at time $t = 0$ to a circuit consisting of a resistance R in series with a capacitance C, there being no initial potential difference across the capacitor.

7 Sequences and series

Summary

This chapter introduces the idea of sequences, such concepts proving particularly relevant in considerations of digital signals which can be thought of as sequences of pulses. The main aspect of the chapter is, however, series and the use of the Fourier series to represent non-sinusoidal signals.

Objectives

By the end of this chapter, the reader should be able to:

- understand what is meant by a sequence and uses the idea to describe digital signals;
- recognise arithmetic and geometric series;
- recognise that some series can converge to a limit, determining the sums of such series;
- recognise the binomial series and uses it in engineering problems;
- represents waveforms by Fourier series and applies the series in the analysis of electrical circuit problems involving non-sinusoidal signals.

7.1 Sequences and series

This section considers what is meant by sets and sequences, considering some commonly encountered forms and their relevance to engineering.

7.1.1 Sequences

Consider the numbers, 1, 3, 5, 7, 9. Such a set of numbers is termed a *sequence* because the numbers are stated in a definite order, 1 followed by 3 followed by 5, etc. Another sequence might be $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$. These sequences have a finite number of terms but often we can meet ones involving an infinite number of terms, e.g. 2, 4, 6, 8, 10, 12, ..., etc.

Key point

A sequence is a set of quantities stated in a definite order.

The term sequence is used for a set of quantities stated in a definite order.

In general we can write a sequence as:

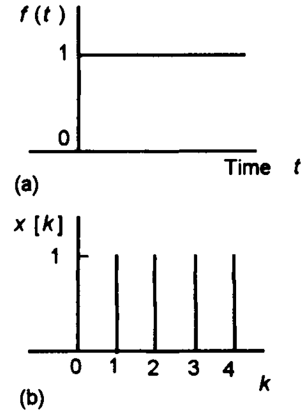


Figure 7.1 *(a) Unit step, (b) unit step sequence*

first value of variable, second value of variable, third value of variable, ..., etc.

or, if x is the variable:

$x[1]$, $x[2]$, $x[3]$, ..., etc.

This is usually more compactly written as $x[k]$, where k = 1, 2, 3, ..., etc. Such a form of notation is commonly encountered in signal processing when perhaps an analogue signal is sampled at a number of sequential points and the resulting sequence of digital signal values processed. For example, if an analogue unit step signal is sampled the sampled data output might be expressed as $x[k] = 0$ for $k < 0$, $x[k] = 1$ for $k \geq 0$ with k = 0, 1, 2, 3, 4, etc. Figure 7.1 shows graphs of the unit step input and the sampled output.

Sometimes it is possible to describe a sequence by giving a rule for the kth term, common forms being the arithmetic and geometric sequences.

- ***Arithmetic sequence***
  An arithmetic sequence has each term formed from the previous term by simply adding on a constant value. If a is the first term and d the common difference between successive terms, the terms are:

  a, $(a + d)$, $(a + 2d)$, $(a + 3d)$, ..., etc. [1]

  The kth term is $a + (k - 1)d$, with k = 1, 2, 3, 4, ..., etc. (note that if k has the values 0, 1, 2, etc. the kth term is $a + kd$). Thus for such a sequence we can write:

  $x[k] = a + (k - 1)d$ [2]

- ***Geometric sequence***
  A geometric sequence has each term formed from the previous term by multiplying it by a constant factor, e.g. 3, 6, 12, 24, ... If a is the first term and r the common ratio between successive terms, the terms are:

  a, ar, ar^2, ar^3 + ..., etc. [3]

  The kth term is ar^{k-1}, with k = 1, 2, 3, 4, ..., etc. Thus for such a sequence we can write:

  $x[k] = ar^{k-1}$ [4]

- ***Harmonic sequence***
  The sequence 1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, ... is termed the *harmonic sequence* and defined for k = 1, 2, 3, etc. by:

Key points

An arithmetic sequence has each term formed from the previous term by simply adding on a constant value.

$x[k] = a + (k - 1)d$

A geometric sequence has each term formed from the previous term by multiplying it by a constant factor, e.g. 3, 6, 12, 24, ...

$x[k] = ar^{k-1}$

$$x[k] = \frac{1}{k} \quad [5]$$

Sequences can be generated by other rules. For example, the sequence 1, 2, 5, 10, 17, ... is generated by $x[k] = 1 + (k - 1)^2$, where $k = 1, 2, 3, \ldots$. This sequence is neither an arithmetic nor a geometric sequence.

Example

Write down the first five terms of the sequence $x[k]$ defined by $x[k] = \frac{1}{2}k^2 + k$ when $k \geq 0$.

When $k = 0$ we have 0 + 0, when $k = 1$ we have 0.5 + 1, when $k = 2$ we have 2 + 2, and so on. The sequence is thus 0, 1.5, 4, 7.5, 12.

7.1.2 Series

A *series* is formed by adding the terms of a sequence. Thus 1 + 3 + 5 + 7 + 9 + ..., etc. is a series.

Key point

A series is the sum of the terms of a sequence.

A series is the sum of the terms of a sequence.

The sum of n terms of a series is written using *sigma notation* as:

$$S_n = \sum_{k=1}^{n} x[k] \quad [6]$$

The first and the last values of k are shown below and above the sigma. For example, the series 1 + 3 + 5 + 7 + 9 would have the sum, over the five terms, written as:

$$S_5 = \sum_{k=1}^{5} (2k - 1)$$

Common series are:

- ***Arithmetic series***

  An arithmetic series has each term formed from the previous term by simply adding on a constant value. Such a series can be written in the general form as:

  $$a + \{a + d\} + \{a + 2d\} + \{a + 3d\} + \ldots + \{a + (n - 1)d\} \quad [7]$$

  The sum to k terms is:

  $$S_k = \{a\} + \{(a + d)\} + \{(a + 2d)\} + \{(a + 3d)\} + \ldots + \{a + (n - 1)d\}$$

  If we write this back to front then:

$S_k = \{a + (n - 1)d\} + \{a + (n - 2)d\} + \{a + (n - 3)d\} + \dots \{a\}$

Adding these two equations gives first term plus first term, second term plus second term, etc. and we obtain:

$2S_k = \{2a + (n - 1)d\} + \{2a + (n - 1)d\} + \{2a + (n - 1)d\} + ..$

for k terms. Thus $2S_k = n\{2a + (n - 1)d\}$ and so:

$$S_k = \tfrac{1}{2}n\{2a + (n - 1)d\} \qquad [8]$$

- ***Geometric series***
  A geometric series has each term formed from the previous term by multiplying it by a constant factor. Such a series can be written in the general form as:

$$a + ar + ar^2 + ar^3 + \dots + ar^{n-1} \qquad [9]$$

The sum to the kth terms is:

$S_k = a + ar + ar^2 + ar^3 + \dots + ar^{n-1}$

Multiplying by r gives:

$rS_k = ar + ar^2 + ar^3 + ar^4 + \dots + ar^n$

Hence $S_k - rS_k = a - ar^n$, and so, provided $k \neq 1$:

$$S_k = \frac{a(1 - r^n)}{1 - r} \qquad [10]$$

Example

Determine the sum of the arithmetic series 1 + 5 + 9 + ... if it contains 10 terms.

Such a series has a first term a of 1 and a common difference d of 4. Thus, using equation [8]:

$S_k = \tfrac{1}{2}\{2a + (k - 1)d\} = \tfrac{1}{2} \times 10\{2 + 9 \times 4\} = 190$

Example

Determine the sum of the geometric series 4 + 6 + 9 + ... if it contains 10 terms.

Such a series has a first term of 4 and a common ratio of 3/2. Thus, using equation [10]:

$$S_k = \frac{a(1-r^k)}{1-r} = \frac{4(1-1.5^{10})}{1-1.5} = 453.3$$

Convergent and divergent series

So far we have considered the sums of series with a finite number of terms. What about the sum when we have a series with an infinite number of terms?

Key point

A series in which the sum of the series tends to a definite value as the number of terms tends to infinity is called a convergent series.

A series in which the sum of the series tends to a definite value as the number of terms tends to infinity is called a convergent series.

Consider an *arithmetic series* $a + (a + d) + (a + 2d) + \dots$ for an infinite number of terms. For k terms we have the sum (equation [8]) of:

$$S_k = \tfrac{1}{2}n\{2a + (n-1)d\}$$

As k tends to infinity then n tends to infinity and so the sum tends to infinity. The sum of an infinite arithmetic series is infinity. The series is said to be *divergent*.

Consider a *geometric series* $a + ar + ar^2 + \dots$ for an infinite number of terms. For k terms we have the sum (equation [10]) of:

$$S_k = \frac{a(1-r^n)}{1-r} = \frac{a}{1-r} - \frac{ar^n}{1-r} \qquad [11]$$

Suppose we have $-1 < r < 1$, as n tends to infinity then r^n tends to 0. Thus the second term converges to zero and we are left with just the first term. Thus such a series converges to the sum:

$$S_\infty = \frac{a}{1-r} \text{ for } -1 < r < 1 \qquad [12]$$

Thus the geometric series $x[k] = 3^{1/2}$ converges to the sum 6. However, if we had the geometric series $x[k] = 3^2$ then the sum is given by equation [15] as $-3 + 3 \times 2^n$ and thus as n tends to infinity the sum tends to infinity. For $|r| \geq 1$ the geometric series does not converge.

There are a number of ways that are used to determine whether a series will converge:

- ***Comparison test***
  A series of positive terms is convergent if its terms are less than the corresponding terms of a positive series which is known to converge. Similarly, the series is divergent if its terms are greater than the corresponding terms of a series which is known to be divergent. As an example, consider:

  $$1 + \frac{1}{2^2} + \frac{1}{3^3} + \frac{1}{4^4} + \dots$$

Suppose we know that the series:

$$1+\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\dots$$

converges (it is a geometric series with r = ½), then if, after the first two terms, we compare terms we find that every term in our convergent series is greater than the one we are considering. Thus the series must also converge.

Key point

The size of a real number x is called its modulus and denoted by $|x|$.

- ***D'Alembert's ratio test***
  An infinite series is convergent if, as k tends to infinity, the ratio of each term u_{n+1} to the preceding term u_n is numerically less than 1 and divergent if greater than 1, i.e.

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| < 1, \text{ the series converges,}$$

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| > 1, \text{ the series diverges,}$$

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = 1, \text{ the series may converge or diverge.}$$

Key point

The symbol ! appearing after a number means that it is multiplied by all the integers between it and 1, e.g. 5! = 5 × 4 × 3 × 2 × 1.

Consider the series:

$$1-\frac{1}{2!}+\frac{1}{3!}-\frac{1}{4!}+\dots$$

The nth term u_n is $|1/n!|$ and the $(n + 1)$th term u_{n+1} is $|1/(n + 1)!|$. Therefore:

$$\frac{u_{n+1}}{u_n}=\left|\frac{\frac{1}{(n+1)!}}{\frac{1}{n!}}\right|=\left|\frac{1}{n+1}\right|$$

As n tends to infinity then:

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| < 1$$

and so the series converges.

Example

Find the sum to infinity of the series 4 + 2 + 1 + ½ + ...

This is a geometric series with a = 4 and r = ½. Using equation [12]:

$$S_\infty=\frac{a}{1-r}=\frac{4}{1-\frac{1}{2}}=8$$

Example

Determine, using the comparison test, whether the series $x[k] = 1/n^n$, i.e. $1 + 1/2^2 + 1/3^3 + 1/4^4 + \dots$, is convergent.

If we exclude the first two terms we can compare it with the geometric series $1/2^3 + 1/2^4 + 1/2^5 + \dots$ which is known to be convergent. Each term in this series being tested is smaller than the comparable term in the comparison series. Thus it must be convergent.

Example

Determine, using d'Alembert's ratio test, whether the series $1 + x + x^2/2! + x^3/3! + \dots$ is convergent.

Using d'Alembert's ratio test, since $u_n = x^{n-1}/(n-1)!$ and $u_{n+1} = x^n/n!$:

$$\frac{u_{n+1}}{u_n} = \frac{\dfrac{x^{n-1}}{(n-1)!}}{\dfrac{x^n}{n!}} = \frac{x}{n}$$

In the limit as n tends to infinity then the ratio tends to 0. Thus the series is convergent.

Power series

A series of the type:

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_nx^n + \dots$$

is known as a *power series*. If we apply d'Alembert's ratio test then the series will be convergent when:

$$\lim_{n\to\infty}\left|\frac{a_{n+1}x^{n+1}}{a_nx^n}\right| < 1$$

This can be written as:

$$|x|\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| < 1$$

or:

$$|x| < \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| \qquad [13]$$

Thus there are conditions attached to the value of x if the series is to converge. Examples are given later in this chapter.

Example

For what values of x is the series $x[k] = x^n/n$ convergent?

Here $a_n = 1/n$ and $a_{n+1} = 1/(n + 1)$. Thus $|a_{n+1}/a_n| = (n + 1)/n = 1 + 1/n$ and so in the limit we have the value of 1 for the limit. Thus the condition for convergence is that $|x| < 1$ or $-1 < x < +1$.

Binomial series

For $(1 + x)^2$ we can readily show that it can be written as $1 + 2x + x^2$. If we multiply this by $(1 + x)$ we obtain $1 + 3x + 3x^2 + x^3$. Multiplying by repeated factors of $(1 + x)$ enables expansions of higher powers of $(1 + x)$ to be generated. This is, however, rather cumbersome if, say, we wanted the expansion of $(1 + x)^{10}$. There is, however, a pattern in the results:

$$\begin{aligned}
(1+x)^1 &= 1 + 1x \\
(1+x)^2 &= 1 + 2x + 1x^2 \\
(1+x)^3 &= 1 + 3x + 3x^2 + 1x^3 \\
(1+x)^4 &= 1 + 4x + 6x^2 + 4x^3 + 1x^4
\end{aligned}$$

If we just write the coefficients the pattern is more readily discerned:

```
      1   1
    1   2   1
  1   3   3   1
1   4   6   4   1
```

Every coefficient is obtained by adding the two either side of it in the row above. Thus, for example, we have:

```
1     1          1     3
   2                4
```

The above pattern is known as *Pascal's triangle*.

However, we can show that the above pattern can be given by:

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \qquad [14]$$

This is known as the *Binomial theorem*. The theorem can be used for both positive and negative values of n and fractional values. With n a positive number the series will eventually terminate. With n a negative number, the series does not terminate. The series converges if we have $-1 < x < 1$.

Example

Expand by the binomial theorem $(1 + x)^6$.

$$(1+x)^6 = 1 + 6x + \frac{6\times5}{2!}x^2 + \frac{6\times5\times4}{3!}x^3 + \frac{6\times5\times4\times3}{4!}x$$

$$+\frac{6\times5\times4\times3\times2}{5!}x^5 + \frac{6\times5\times4\times3\times2\times1}{6!}x^6$$

$$= 1 + 6x + 15x^2 + 20x^3 + 15x^4 + 6x^5 + x^6$$

Example

Write the first four terms in the expansion of $(1 + x)^{1/2}$.

$$(1+x)^{1/2} = 1 + \tfrac{1}{2}x + \frac{\frac{1}{2}\left(-\frac{1}{2}\right)}{2!}x^2 + \frac{\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{3!}x^3 + \ldots$$

$$= 1 + \tfrac{1}{2}x - \tfrac{1}{8}x^2 + \tfrac{1}{16}x^3 + \ldots$$

Maths in action

Making approximations

A common use of the Binomial theorem in engineering is for making approximations. For example, we might want to determine the change in the second moment of area of a rectangle which was given by $bL^3/12$ if b is increased by 3% and L reduced by 2%. The new second moment of area is:

$$I = \tfrac{1}{12}(1+0.03)b\left[(1-0.02)L\right]^3$$

$$= \frac{bL^3}{12}(1+0.03)(1-0.02)^3$$

Using the Binomial theorem for the cubed term and neglecting, since they will be very small, all terms involving powers of 0.02:

$$I = \frac{bL^3}{12}(1+0.03)(1-0.06+..) = \frac{bL^3}{12}(1+0.03-0.06+..)$$

Hence, $I = 0.97bL^3/12$ and so the percentage change is a reduction by approximately 3%.

Useful power series

Table 7.1 gives some commonly met functions and their series expansions.

Table 7.1 *Power series*

Function	Series	Validity
$\sin x$	$x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots$	For all x
$\cos x$	$x-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots$	For all x
$\tan x$	$x+\frac{x^3}{3!}+\frac{2x^5}{15}+\frac{17x^7}{315}+\ldots$	$-\pi/2 < x < \pi/2$
e^x	$1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\ldots$	For all x
$\sinh x$	$x+\frac{x^3}{3!}+\frac{x^5}{5!}+\frac{x^7}{7!}+\ldots$	For all x
$\cosh x$	$1+\frac{x^2}{2!}+\frac{x^4}{4!}+\frac{x^6}{6!}+\ldots$	For all x
$\ln(1+x)$	$x-\frac{x^2}{2!}+\frac{x^3}{3!}-\frac{x^4}{4!}+\ldots$	$-1 < x < 1$

Example

Using series given in Table 7.1, determine the series expansion of the function $e^x \sin x$.

Table 7.1 gives:

$\sin x = x-\frac{x^3}{3!}+\frac{x^5}{5!}+\ldots$, valid for all values of x

$e^x = 1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots$, valid for all values of x

We can multiply these two series to give

$$e^x \sin x = \left(1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots\right)\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}+\ldots\right)$$

$$= x + x^2 + \left(\tfrac{1}{2}-\tfrac{1}{6}\right)x^3+\left(\tfrac{1}{6}-\tfrac{1}{6}\right)x^4 + \left(\tfrac{1}{120}+\tfrac{1}{24}-\tfrac{1}{12}\right)x^5+\ldots$$

$$= x + x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{30}x^5+\ldots$$

Example

Using Table 6.1, determine the series, as far as the x^3 term, for the function $y = e^{4x}$.

Table 6.1 gives:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

If we substitute $4x$ for x then we obtain:

$$e^{4x} = 1 + \frac{4x}{1!} + \frac{16x^2}{2!} + \frac{64x^3}{3!} + \dots = 1 + 4x + 8x^2 + \frac{32}{3}x^3 + \dots$$

Problems 7.1

1 A sinusoidal signal $f(t) = \sin t$ is sampled every quarter period starting when $t = 0$. State the sequence of sampled values.

2 Write down the first five terms of the sequence $x[k]$ defined, for $k \geq 0$, by (a) $x[k] = k$, (b) $x[k] = e^{-k}$.

3 State the fifth term of (a) the arithmetic sequence given by 4, 7, 10, ..., (b) the geometric sequence given by 12, 6, 3, ...

4 Write an equation for the kth term, where $k = 1, 2, 3, \dots$, for the following sequences (a) 1, –1, 1, –1, ..., (b) 5, 10, 15, 20, ..., (c) 2, 1, 5, 1, 0.5, ...

5 Write down the first five terms of the sequence $x[k]$ defined, for $k \geq 0$, by (a) $x[k] = k^2$, (b) $x[k] = e^k$, (c) $x[k] = \frac{1}{2}k^2 + 2k$.

6 State the fifth term of the arithmetic progression given by 5, 7, 9, ...

7 State the fifth term of the geometric progression given by 8, 4, 2, ...

8 Write an equation for the kth term for the following sequences: (a) $\frac{1}{4}, \frac{1}{16}, \frac{1}{64}, \dots$, (b) –2, +2, –2, ..., (c) 3.1, 3.01, 3.001, ...

9 State the first three terms of the sequences given by:

(a) $(0.1)^k$, (b) $5 + (0.1)^k$, (c) $(-1)^k$

10 Determine the sums of the following series if each contains 12 terms:

(a) 2 + 5 + 8 + ..., (b) $5 + \frac{5}{2} + \frac{5}{4} + \dots$,
(c) 4 + 3.6 + 3.24 + ...

11 Determine the sums of the following arithmetic or geometric series if each contains 10 terms:

(a) 3 + 2.5 + 2.0 + ..., (b) 12 + 6 + 3 + ...,
(c) 1 + 2 + 4 + 8 + ...

12 Find the sum to infinity of the series:

(a) 6 + 3 + 1.5 + ..., (b) 4 + 3 + 2.25 + ...,
(c) 12 + 3 + 0.75 + ...

13 Using the comparison test, determine whether the following series are convergent or divergent:

(a) $x[k] = 1/3^n$ (compare with $1/2^n$), (b) $x[k] = 1.5^n$ (compare with 1^n)

14 Using d'Alembert's ratio test, determine whether the following series are convergent or divergent:

(a) $x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots + \frac{(-1)^{n-1}x^n}{n} + \ldots$

(b) $3 + \frac{3^2}{2} + \frac{3^3}{3} + \ldots + \frac{3^n}{n} + \ldots$

15 Determine which of the following series is convergent and which divergent:

(a) $-1 + 1 - 1 + \ldots (-1)^n + \ldots$,

(b) $1\,e^{-1} + 2\,e^{-2} + 3\,e^{-3} + \ldots n\,e^{-n} + \ldots$,

(c) $\frac{2\times1+1}{2} + \frac{2\times2+1}{2^2} + \frac{2\times3+1}{2^3} + \ldots + \frac{2n+1}{2^n} + \ldots$,

(d) $\sum_{k=0}^{\infty} \frac{10^n}{n!}$, (e) $\sum_{k=1}^{\infty} \frac{1}{n^2+2^2}$, (f) $\sum_{k=1}^{\infty} \frac{2^{n-1}}{10+(n-1)}$

16 Expand by the binomial theorem:

(a) $(1 + x)^4$, (b) $(1 + x)^{3/2}$, (c) $(1 - x)^{-5/2}$,

(d) $(1 + 0.25)^{-1}$ for four terms, (e) $(4 + x)^{1/2}$ for four terms.

17 Use the binomial theorem to write the first four terms of:

(a) $(1 + x)^{12}$, (b) $(1 - 2x)^{-2}$, (c) $(3 - 2x)^{2/5}$, (d) $\frac{1}{1-x}$,

(e) $(1 + 3x)^{-1/2}$, (f) $\frac{1}{(1-x^3)^2}$

18 By using the binomial theorem, determine the cube root of 1.04 to four decimal places. Hint: write 1.04 as 1 + 0.04.

19 The transverse deflection δ of a column of length L when subject to a vertical load F and a horizontal load H at the top is given by:

$$\delta = \frac{HL}{F}\left(\frac{\tan aL}{aL} - 1\right)$$

where $a^2 = F/EI$. Show that as F tends to a zero value that δ tends to $HL^3/3EI$.

20 Determine the series expansion for cosh x using the relationship $\cosh x = \frac{1}{2}(e^x + e^{-x})$.

21 Determine the series expansion for tan x using $\tan x = \sin x/\cos x$.

22 Using Table 7.1, determine the series for the following functions:

(a) $y = e^{2x}$, (b) $y = e^x \cos x$, (c) $y = (1 + x)^{-1/2}$,

(d) $y = e^x \ln(1 + x)$, (e) $y = \sec x$, (f) $y = \cos^2 x$

23 Show that, if x is small:

$$\frac{1}{1+x} - (1-2x)^{1/2} \approx \tfrac{3}{2}x^2$$

24 For a continuous belt passing round two wheels, diameters d and D, with centres a distance x apart, the length L of belt required, if there is no sag, is:

$$L = 2x\cos\alpha + \tfrac{1}{2}\pi(D+d) + (D-d)\alpha$$

where $\sin\alpha = (D-d)/2x$. Show that:

$$L \approx 2x + \tfrac{1}{2}\pi(D+d) + \frac{(D-d)^2}{4x}$$

25 The displacement x of the slider of a reciprocating mechanism depends on the crankshaft angle θ, being related by

$$x = r\cos\theta + L\sqrt{1 - \frac{r^2}{L^2}\sin^2\theta}$$

where r is the radius of the crankshaft and L the length of the connecting rod. Show, when r/L is considerably smaller than 1, that:

$$x \approx r\cos\theta + L - \frac{r^2}{2L}\sin^2\theta$$

26 Determine the approximate percentage change in the volume of a cylinder if its radius is reduced by 4% and its height increased by 2%.

27 The resonant frequency of an electrical circuit containing capacitance C and inductance L is given by $1/[2\pi\sqrt{(LC)}]$. Determine the approximate percentage change in the frequency if the capacitance is increased by 2% and the inductance decreased by 1%.

7.2 Fourier series

Alternating waveforms in, say, electrical circuits are not always sinusoidal. For example, many voltages which might initially have been sinusoidal have their waveforms 'distorted' by being applied to some non-linear device and thus we need to be able to considered the behaviour of such a waveform with an electrical circuit. In other cases we might have a rectangular waveform rather than a sinusoidal one. This section is a consideration of how we can use a series to describe such waveforms.

7.2.1 Fourier series

Key point

Fourier series:
Any periodic waveform can be represented by a constant d.c. signal term plus terms involving sinusoidal waveforms of multiples of a basic frequency.

The Fourier series is concisely expressed as:

$$y = A_0 + \sum_{n=1}^{\infty} A_n \sin\left(n\omega t + \phi_n\right)$$

In 1822 Jean Baptiste Fourier proposed that any periodic waveform could be made up of a combination of sinusoidal waveforms, i.e.

$$y = A_0 + A_1 \sin(1\omega t + \phi_1) + A_2 \sin(2\omega t + \phi_2) + A_3 \sin(3\omega t + \phi_3) + \ldots \quad [15]$$

This is termed the *Fourier* series. where A_0 is a non-alternating component, e.g. a d.c. component. The waveform element with the $1\omega t$ frequency is called the *fundamental frequency* or the *first harmonic*, the element with the $2\omega t$ frequency the *second harmonic*, the element with the $3\omega t$ the third harmonic, and so on. A_1, A_2, A_3 are the amplitudes of the components and ϕ_1, ϕ_2, ϕ_3, etc. their phases.

As an illustration, consider the waveform produced by having just sine terms with the fundamental and the third harmonic and $A_3 = A_1/3$, i.e.

$$y = A_1 \sin 1\omega t + \tfrac{1}{3}A_1 \sin 3\omega t$$

Figure 7.2 shows graphs of the two terms and the waveform obtained by adding the two, ordinate by ordinate.

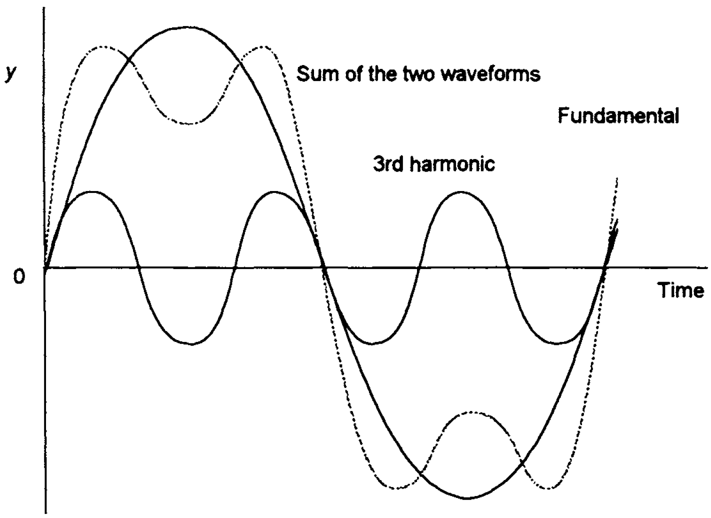


Figure 7.2 *Adding two waveforms*

The result of adding the two waveforms is something that begins to look a bit like a rectangular waveform. The addition of a d.c. term shifts the waveform up or down. If we add a d.c. term of $0.79\,A_1$ then Figure 7.2 becomes transformed to Figure 7.3:

$$y = 0.79A_1 + A_1 \sin 1\omega t + \tfrac{1}{3}A_1 \sin 3\omega t$$

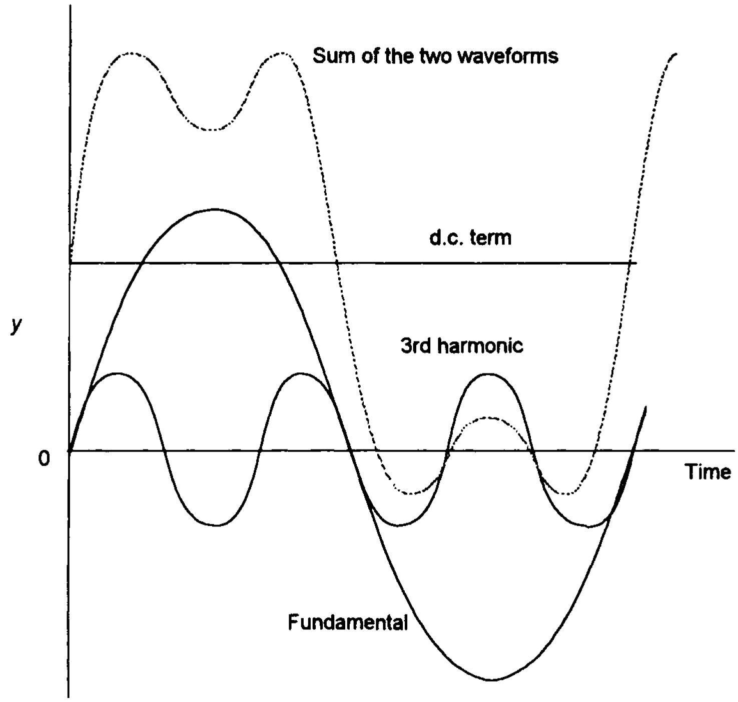


Figure 7.3 *Adding a d.c. term*

A better approximation to a rectangular waveform is given by adding more terms:

$$y = 0.79A_1 + A_1 \sin 1\omega t + \tfrac{1}{3}A_1 \sin 3\omega t + \tfrac{1}{5}A_1 \sin 5\omega t + \tfrac{1}{7}A_1 \sin 7\omega t + \ldots$$

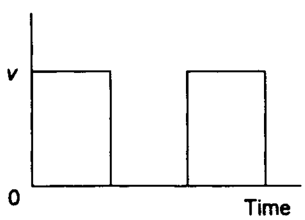


Figure 7.4 *Rectangular waveform*

We then obtain a rectangular waveform which approximates to a periodic sequence of pulses (Figure 7.4).

Alternative way of writing the Fourier series

There is an alternative, simpler, way of writing equation [15]. Since $\sin(A + B) = \sin A \cos B + \cos A \sin B$ we can write:

$$A_1 \sin(1\omega t + \phi_1) = A_1 \sin \phi_1 \cos 1\omega t + A_1 \cos \phi_1 \sin 1\omega t$$

If we represent the non-time varying terms $A_1 \sin \phi_1$ by a constant a_1 and $A_1 \cos \phi_1$ by b_1, then:

$A_1 \sin(1\omega t + \phi_1) = a_1 \cos 1\omega t + b_1 \sin 1\omega t$

Likewise we can write:

$A_2 \sin(2\omega t + \phi_2) = a_2 \cos 2\omega t + b_2 \sin 2\omega t$

$A_3 \sin(3\omega t + \phi_3) = a_3 \cos 3\omega t + b_3 \sin 3\omega t$

Key point

Note: in Figure 7.3 the addition of a d.c. term of 0.79 to the waveform results in an average value of this waveform over one cycle of 0.79. The term $\frac{1}{2}a_0$ in the Fourier series thus represents the average value of the waveform over a cycle.

and so on. If, for convenience we choose to write $\frac{1}{2}a_0$ for A_0, equation [10] can be written as:

$$y = \tfrac{1}{2}a_0 + a_1 \cos 1\omega t + a_2 \cos 2\omega t + a_3 \cos 3\omega t + \dots \\ + b_1 \sin 1\omega t + b_2 \sin 2\omega t + b_3 \sin 3\omega t + \dots \qquad [16]$$

Hence we can write the Fourier series equation as:

$$y = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega t + \sum_{n=1}^{\infty} b_n \sin n\omega t \qquad [17]$$

The a and b terms are called the *Fourier coefficients*.
Since we have $a_n = A_n \sin \phi_n$ and $b_n = A_n \cos \phi_n$ then:

$$\phi_n = \tan^{-1}\left(\frac{a_n}{b_n}\right) \qquad [18]$$

and, since:

$$a_n^2 + b_n^2 = A_n^2 \sin^2\phi_n + A_n^2 \cos^2\phi_n = A_n^2$$

we have:

$$A_n = \sqrt{a_n^2 + b_n^2} \qquad [19]$$

7.2.2 Fourier coefficients

Now consider how we can establish the Fourier coefficients for a waveform. Suppose we have the Fourier series in the form of equation [16]:

$$y = \frac{1}{2}a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + \dots + a_n \cos n\omega t \\ + b_1 \sin \omega t + b_2 \sin 2\omega t + \dots + b_n \sin n\omega t$$

If we integrate both sides of the equation over one period T of the fundamental, the integral for each cosine and sine term will be the area under the graph of that expression for one cycle and thus zero. A consequence of this is that the only term which is not zero when we integrate the equation is the integral of the a_0 term. Thus, integrating over one period T gives:

$$\int_0^T y \, dt = \int_0^T \frac{1}{2}a_0 \, dt = \frac{1}{2}a_0 T$$

and so:

$$a_0 = \frac{2}{T}\int_0^T y\,dt \qquad [20]$$

We can obtain the a_1 term by multiplying the equation by cos ωt and then integrating over one period. Thus the equation becomes:

$$y\cos\omega t = \tfrac{1}{2}a_0\cos\omega t + a_1\cos\omega t\cos\omega t + a_2\cos\omega t\cos 2\omega t + \dots + b_1\cos\omega t\sin\omega t + b_2\cos\omega t\sin 2\omega t + \dots$$

$$= \tfrac{1}{2}a_0\cos\omega t + a_1\cos^2\omega t + a_2\cos\omega t\cos 2\omega t + \dots + b_1\cos\omega t\sin\omega t + b_2\cos\omega t\sin 2\omega t + \dots$$

The integration over a period T of all the terms involving sin ωt and cos ωt will be zero. Thus we are only left with the $\cos^2\omega t$ term and so, using equation [13]:

$$\int_0^T y\cos\omega t\,dt = \int_0^T a_1\cos^2\omega t\,dt$$

$$= \tfrac{1}{2}a_1\int_0^T(1+\cos 2\omega t)\,dt$$

$$= \tfrac{1}{2}a_1\left[t + \frac{\sin 2\omega t}{2\omega}\right]_0^T = \tfrac{1}{2}a_1 T$$

and so we have:

$$a_1 = \frac{2}{T}\int_0^T y\cos\omega t\,dt \qquad [21]$$

In general, multiplying the equation by cos $n\omega t$ gives:

$$a_n = \frac{2}{T}\int_0^T y\cos n\omega t\,dt \qquad [22]$$

This equation gives for $n = 0$ the equation given earlier for a_0. This would not have been the case if the first term in the Fourier series had been written as a_0 instead of $a_0/2$.

In a similar way, multiplying the equation by sin ωt and integrating over a period enables us to obtain the b coefficients. Thus:

$$y\sin\omega t = \tfrac{1}{2}a_0\sin\omega t + a_1\sin\omega t\cos\omega t + a_2\sin\omega t\cos 2\omega t + \dots + b_1\sin\omega t\sin\omega t + b_2\sin\omega t\sin 2\omega t + \dots$$

$$= \tfrac{1}{2}a_0\sin\omega t + a_1\sin\omega t\cos\omega t + a_2\sin\omega t\cos 2\omega t + \dots + b_1\sin^2\omega t + b_2\sin\omega t\sin 2\omega t + \dots$$

The integration over a period T of all the terms involving sin ωt and cos ωt will be zero and so:

$$\int_0^T y\sin\omega t\,dt = \int_0^T b_1\sin^2\omega t\,dt$$

$$= \frac{1}{2}\int_0^T b_1(1 - \cos 2\omega t)\,dt$$

$$= \frac{1}{2}b_1\left[t - \frac{\sin 2\omega t}{2\omega}\right]_0^T = \frac{1}{2}b_1 T$$

Hence:

$$b_1 = \frac{2}{T}\int_0^T y \sin \omega t\,dt \qquad [23]$$

In general, multiplying the equation by sin $n\omega t$ and integrating gives:

$$b_n = \frac{2}{T}\int_0^T y \sin n\omega t\,dt \qquad [24]$$

The following illustrates how the Fourier series can be established for a number of common waveforms.

Rectangular waveform

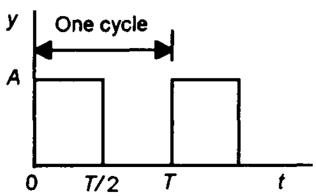


Figure 7.4 *Rectangular waveform*

Consider the *rectangular waveform* shown in Figure 7.4. It can be described as:

$y = A$ for $0 \le t < T/2$, and $y = 0$ for $T/2 \le t < T$, period T

Now consider the determination of the coefficients. Equation [20] for a_0:

$$a_0 = \frac{2}{T}\int_0^T y\,dt$$

has an integral which is the area under the graph of y against t for the period T. Since this area is $AT/2$, we have $a_0 = A$. To obtain a_n we use equation [22]:

$$a_n = \frac{2}{T}\int_0^T y \cos n\omega t\,dt$$

Since y has the value A up to $T/2$ and is zero thereafter, we can write the above equation in two parts as:

$$a_n = \frac{2}{T}\int_0^{T/2} A \cos n\omega t\,dt + \frac{2}{T}\int_{T/2}^T 0 \cos n\omega t\,dt$$

The value of the second integral is 0 and so:

$$a_n = \frac{2}{T}\left[\frac{A}{n\omega}\sin n\omega t\right]_0^{T/2}$$

Since $\omega = 2\pi/T$ then the sine term is sin $2n\pi t/T$. Thus with $t = T/2$ we have sin $n\pi$ which is zero and since sin 0 = 0, we have $a_n = 0$.

For the b_n terms we use equation [24]:

$$b_n = \frac{2}{T}\int_0^T y \sin n\omega t\,dt$$

Since we have $y = A$ from 0 to $T/2$ and then $y = 0$ for the remainder of the period, this equation can be written in two parts as:

$$b_n = \frac{2}{T}\int_0^{T/2} A \sin n\omega t \, dt + \frac{2}{T}\int_{T/2}^{T} 0 \sin n\omega t \, dt$$

The value of the second integral is 0 and so:

$$b_n = \frac{2}{T}\left[-\frac{A}{n\omega}\cos n\omega t\right]_0^{T/2} = \frac{A}{\pi n}(1 - \cos n\pi)$$

Hence:

$$b_1 = \frac{A}{\pi}(1 - \cos\pi) = \frac{2A}{\pi}, \quad b_2 = \frac{A}{2\pi}(1 - \cos 2\pi) = 0$$

$$b_3 = \frac{A}{3\pi}(1 - \cos 3\pi) = \frac{2A}{3\pi}, \quad \text{etc.}$$

Thus the Fourier series for the rectangular waveform can be written as:

$$y = A\left(\frac{1}{2} + \frac{2}{\pi}\sin\omega t + \frac{2}{3\pi}\sin 3\omega t + \dots\right) \qquad [25]$$

Note that only odd harmonics are present.

Sawtooth waveform

Consider the *sawtooth waveform* shown in Figure 7.5. It can be described by:

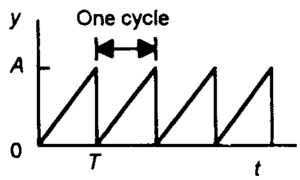


Figure 7.5 *Sawtooth waveform*

$y = At/T$ for $0 \le t < T$, period T

To determine a_0 we use equation [20]:

$$a_0 = \frac{2}{T}\int_0^T y \, dt$$

The integral is the area under the graph of y against t between 0 and time T. This is $AT/2$ and so $a_0 = A$. To obtain a_n we use equation [22]:

$$a_n = \frac{2}{T}\int_0^T y \cos n\omega t \, dt$$

Since $\omega = 2\pi/T$ and $y = At/T$ then:

$$a_n = \frac{2}{T}\int_0^T \frac{At}{T}\cos\frac{2\pi nt}{T} \, dt$$

Using integration by parts gives:

$$a_n = \frac{2}{T}\left[\frac{At}{2\pi n}\sin\frac{2\pi nt}{T} + \frac{At}{4\pi^2 n^2}\cos\frac{2\pi nt}{T}\right]_0^T$$

$$= \frac{2A}{T}\left[\frac{T}{4\pi^2 n^2} - \frac{T}{4\pi^2 n^2}\right] = 0$$

The values of a_n are zero for all values other than a_0. The values of b_n can be found by using equation [24]:

$$b_n = \frac{2}{T}\int_0^T y \sin n\omega t \, dt = \frac{2}{T}\int_0^T \frac{At}{T} \sin\frac{2\pi nt}{T} \, dt$$

Integration by parts gives:

$$b_n = \frac{2}{T}\left[-\frac{At}{2\pi n}\cos\frac{2\pi nt}{T} + \frac{At}{4\pi^2 n^2}\sin\frac{2\pi nt}{T}\right]_0^T$$

$$= \frac{2A}{T}\left[-\frac{T}{2\pi n}\right]_0^T = -\frac{A}{\pi n}$$

The Fourier series for the sawtooth waveform is thus:

$$y = \frac{A}{2} - \frac{A}{\pi}\sin\omega t - \frac{A}{2\pi}\sin 2\omega t - \frac{A}{3\pi}\sin 3\omega t - \ldots \qquad [26]$$

We can write this as:

$$y = \frac{A}{2} + \frac{A}{\pi}\cos(\omega t + \frac{\pi}{2}) + \frac{A}{2\pi}\cos(2\omega t + \frac{\pi}{2}) + \frac{A}{3\pi}\cos(3\omega t + \frac{\pi}{2}) + \ldots$$

Half-wave rectified sinusoid

Consider a half-rectified sinusoidal waveform of period T (Figure 7.6). This can be described by:

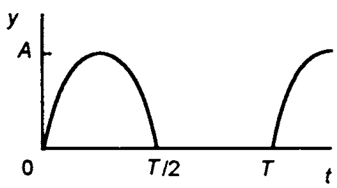


Figure 7.6 *Half-wave rectified sinusoid*

$y = A \sin \omega t = A \sin 2\pi t/T$ for $0 \le t \le T/2$, $y = 0$ for $T/2 \le t < T$

We can determine a_0 by using equation [20]:

$$a_0 = \frac{2}{T}\int_0^T y \, dt = \frac{2}{T}\left(\int_0^{T/2} A \sin\omega t \, dt + \int_{T/2}^T 0 \, dt\right)$$

$$= -\frac{2A}{T\omega}[\cos\omega t]_0^{T/2} = \frac{2A}{\pi}$$

Equation [22] can be used to determine a_n:

$$a_n = \frac{2}{T}\int_0^T y \cos n\omega t \, dt = \frac{2}{T}\left(\int_0^{T/2} A \sin\omega t \cos n\omega t \, dt + \int_{T/2}^T 0 \, dt\right)$$

Since $2 \sin A \cos B = \sin(A + B) + \sin(A - B)$:

$$a_n = \frac{A}{T}\int_0^{T/2}\left[\sin(1+n)\omega t + \sin(1-n)\omega t\right] dt$$

For $n = 1$ we have:

$$a_1 = \frac{A}{T}\int_0^{T/2} \sin(1+1)\omega t \, dt = -\frac{A}{2T\omega}[\cos\omega t]_0^{T/2} = 0$$

For $n > 1$ we have:

$$a_n = \frac{A}{T}\left[-\frac{1}{(1+n)\omega}\cos(1+n)\omega t - \frac{1}{(1-n)\omega}\cos(1-n)\omega t\right]_0^{T/2}$$

For even values of n we have $\cos(1 + n)\pi = -1$ and $\cos(1 - n)\pi = -1$ and so:

$$a_{n\,\text{even}} = \frac{A}{T}\left(\frac{1}{(1+n)\omega} + \frac{1}{(1+n)\omega} + \frac{1}{(1-n)\omega} + \frac{1}{(1-n)\omega}\right.$$

$$= \frac{A}{\pi}\left(\frac{2}{1+n} + \frac{2}{1-n}\right)$$

$$= \frac{2A}{\pi(1-n^2)}$$

For odd values, other than 1, of n we have $\cos(1 + n)\pi = 1$ and $\cos(1 - n)\pi = 1$. This gives:

$$a_{n\,\text{odd}} = \frac{A}{T}\left(-\frac{1}{(1+n)\omega} + \frac{1}{(1+n)\omega} - \frac{1}{(1-n)\omega} + \frac{1}{(1-n)\omega}\right) = 0$$

The values of b_n can be found using equation [24]:

$$b_n = \frac{2}{T}\int_0^T y\sin n\omega t\,dt = \frac{2}{T}\int_0^{T/2} A\sin\omega t\sin n\omega t\,dt$$

Since $2\sin A\sin B = \cos(A - B) - \cos(A + B)$:

$$b_n = \frac{A}{T}\int_0^{T/2}\left[\cos(1-n)\omega t - \cos(1+n)\omega t\right]dt$$

For $n = 1$ we have:

$$b_n = \frac{A}{T}\int_0^{T/2}\left[1 - \cos(1+n)\omega t\right]dt$$

$$= \frac{A}{t}\left[t - \frac{1}{2\omega}\sin 2\omega t\right]_0^{T/2} = \frac{A}{2}$$

For $n > 1$ we have:

$$b_n = \frac{A}{T}\left[\frac{1}{(1-n)\omega}\sin(1-n)\omega t - \frac{1}{(1+n)\omega}\sin(1+n)\omega t\right]_0^{T/2}$$

Since $\sin(1 - n)\pi = 0$ and $\sin(1 + n)\pi = 0$, we have $b_n = 0$ for all values of n other than 1.

The Fourier series for the half-wave rectified sinusoid is thus:

$$y = \frac{A}{\pi} - \frac{2A}{3\pi}\cos 2\omega t - \frac{2A}{15\pi}\cos 4\omega t + \ldots + \frac{A}{2}\sin\omega t \qquad [27]$$

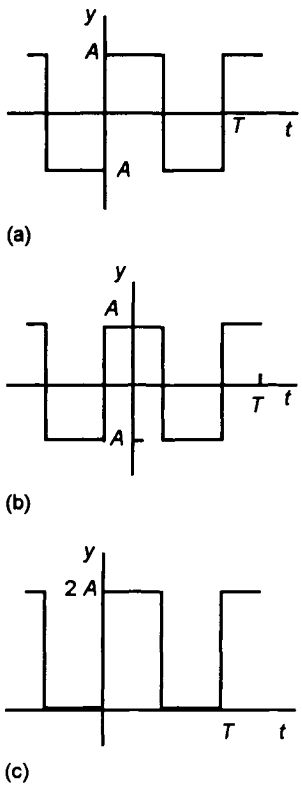


Figure 7.7 *Origin shifts*

Key points

Shifting the time origin of a waveform to the right by θ means replacing t by $(t + \theta)$ in the Fourier series. Shifting the time origin to the left by h means replacing t by $(t - \theta)$.

Shifting the time axis vertically downwards adds to the Fourier series the amount of the shift, shifting upwards subtracts.

Shift of origin

The Fourier series for the rectangular waveform shown in Figure 7.7(a) is:

$$y = \frac{4A}{\pi}\left[\sin\omega t + \tfrac{1}{3}\sin 3\omega t + \tfrac{1}{5}\sin 5\omega t + \dots\right] \qquad [28]$$

Now consider the waveform in Figure 33.7(b). This is the waveform in (a) with the time origin shifted to the right by $\pi/2$. If we work out the Fourier series for this waveform we find that it is equation [28] with t replaced by $(t + \pi/2)$.

$$y = \frac{4A}{\pi}\left[\sin\left(\omega t + \frac{\pi}{2}\right) + \tfrac{1}{3}\sin 3\omega\left(t + \frac{\pi}{2}\right) + \tfrac{1}{5}\sin 5\omega\left(t + \frac{\pi}{2}\right) + \dots\right]$$

and so:

$$y = \frac{4A}{\pi}\left[\cos\omega t - \tfrac{1}{3}\cos 3\omega t + \tfrac{1}{5}\cos 5\omega t - \dots\right] \qquad [29]$$

Thus we have the rule:

> *Shifting the time origin of a waveform to the right by θ means replacing t by $(t + \theta)$ in the Fourier series. Shifting the time origin to the left by h means replacing t by $(t - \theta)$.*

Now consider the waveform in Figure 33.7(c). This is that in (a) shifted vertically by A, i.e. the waveform in (a) plus A. The Fourier series is then that of (a) plus A:

$$y = A + \frac{4A}{\pi}\left[\sin\omega t + \tfrac{1}{3}\sin 3\omega t + \tfrac{1}{5}\sin 5\omega t + \dots\right] \qquad [30]$$

This gives the rule:

> *Shifting the time axis vertically downwards adds to the Fourier series the amount of the shift, shifting upwards subtracts.*

7.2.3 Odd and even symmetry

As will be apparent from the above examples, certain terms are not always present in a Fourier series. Consideration of whether functions have odd or even symmetry about the origin enables us to determine the presence or otherwise of terms.

- ***Odd symmetry***
  A function with odd symmetry is defined as having $f(-t) = -f(t)$. This means that the function value for a particular positive value of time is equal in magnitude but of opposite sign to that for the corresponding negative value of that time. Thus $y = f(x) = x^3$ is an odd function since $f(-2) = -8 = -f(2)$. For every point on the waveform for positive times there is a

corresponding point on the waveform on a straight line drawn through the origin and equidistant from it (Figure 7.8).

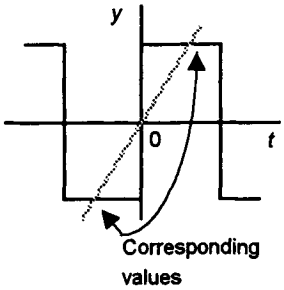


Figure 7.8 *Odd symmetry*

- ***Even symmetry***
  A function with even symmetry is defined as having $f(-t) = f(t)$. This means that the function value for a particular positive value of time is identical to that for the corresponding negative value of that time. Thus $y = f(x) = x^2$ is an even function since $f(-2) = 4 = f(2)$. If the y-axis was a plane mirror then the reflection of the positive time values for the waveform would give the negative time values (Figure 7.9).

In determining Fourier coefficients it is necessary to consider the odd or even nature of products of two odd or even functions.

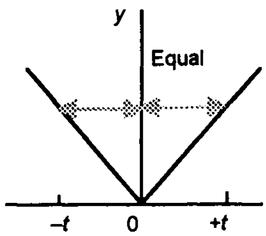


Figure 7.9 *Even symmetry*

- ***Product of two even functions***
  Consider $f(x)$ and $g(x)$ and the product $F(x) = f(x)g(x)$. We can write $F(-x) = f(-x)g(-x)$. Thus if $f(x)$ and $g(x)$ are both even we must have $F(-x) = f(x)g(x)$ and so $F(-x) = F(x)$. The product of two even functions is an even function.

- ***Product of two odd functions***
  Consider $f(x)$ and $g(x)$ and the product $F(x) = f(x)g(x)$. We can write $F(-x) = f(-x)g(-x)$. Thus if $f(x)$ and $g(x)$ are both odd we must have $F(-x) = \{-f(x)\}\{-g(x)\}$ and so $F(-x) = F(x)$. The product of two odd functions is an even function.

- ***Product of an odd and an even function***
  Consider $f(x)$ and $g(x)$ and the product $F(x) = f(x)g(x)$. We can write $F(-x) = f(-x)g(-x)$. Thus if $f(x)$ is even and $g(x)$ is odd we must have $F(-x) = f(x)\{-g(x)\} = -f(x)g(x)$ and so $F(-x) = -F(x)$. The product of an even and an odd function is an odd function.

Example

Determine whether (a) x^2, (b) cos $2x$ and (c) x^2 cos $2x$ are even or odd functions.

(a) $y = f(x) = x^2$ is an even function since if we consider some particular value of x, say –2, we have $f(-2) = 4 = f(2)$.

(b) $y = f(x) = \cos 2x$ is an even function since if we consider some particular value of x, say $-\pi/2$ we have $f(-\pi/2) = 0 = f(\pi/2)$.

(c) Since the product of two even functions is even, $x^2 \cos 2x$ is an even function.

Fourier coefficients for odd/even symmetry

Consider the coefficients for a Fourier series for functions showing odd or even symmetry.

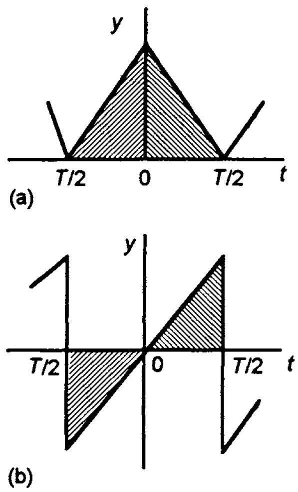


Figure 7.10 *(a) Even, (b) odd*

- ***a_0 coefficients***

  a_0 is given by equation [20] as:

$$a_0 = \frac{2}{T}\int_0^T f(t)\,dt = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\,dt$$

$$= \frac{2}{T}\int_0^{T/2} f(t)\,dt + \frac{2}{T}\int_{-T/2}^{0} f(t)\,dt$$

For a function with even symmetry we have the areas under the waveform on each side of the y-axis equal in both size and sign. Figure 7.10(a) illustrates this. A consequence of this is:

$$a_0 = 2 \times \frac{2}{T}\int_0^{T/2} f(t)\,dt \qquad [31]$$

But for an odd function (Figure 7.10(b)) the areas under the waveform on each side of the y-axis are equal in size but opposite in sign. A consequence of this is that there can be no a_0 term:

$$a_0 = \int_0^{T/2} f(t)\,dt + \int_{-T/2}^{0} f(t)\,dt = 0 \qquad [32]$$

We can look at this issue in another way. The mean value over one cycle of a waveform is $a_0/2$. Thus for an odd function the mean value is 0 because the mean value is 0.

- ***a_n coefficients***

  For the a_n coefficients equation [22] gives:

$$a_n = \frac{2}{T}\int_0^T f(t)\cos n\omega t\,dt = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\cos n\omega t\,dt$$

Since cos $n\omega t$ is an even function, if $f(t)$ is even then the product is even. Hence we have, on the basis of the discussion used for a_0:

$$a_n = 2 \times \frac{2}{T}\int_0^{T/2} f(t)\cos n\omega t\,dt \qquad [33]$$

If $f(t)$ is odd then the product is odd. Thus on the basis of the discussion used for a_0:

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\cos n\omega t\,dt = 0 \qquad [34]$$

- ***b_n coefficients***

  For the b_n coefficients equation [24] gives:

$$b_n = \frac{2}{T}\int_0^T f(t)\sin n\omega t\,dt = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\sin n\omega t\,dt$$

Since sin $n\omega t$ is an odd function, if $f(t)$ is even then the product is odd. Thus, on the basis of the discussion used for a_0:

$$b_n = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\sin n\omega t\,\mathrm{d}t = 0 \qquad [35]$$

If $f(t)$ is odd then the product is even. Thus, on the basis of the discussion used for a_0:

$$b_n = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\sin n\omega t\,\mathrm{d}t = 2\times\frac{2}{T}\int_0^{T/2} f(t)\sin n\omega t\,\mathrm{d}t \qquad [36]$$

Key points

If f(t) is an even function then the Fourier series contains an a_0 term and only cosine terms. If f(t) is an odd function then the Fourier series contains no a_0 term and only sine terms.

To summarise:

If f(t) is an even function then the Fourier series contains an a_0 term and only cosine terms. If f(t) is an odd function then the Fourier series contains no a_0 term and only sine terms.

Example

Determine the Fourier series for the function shown in Figure 7.11.

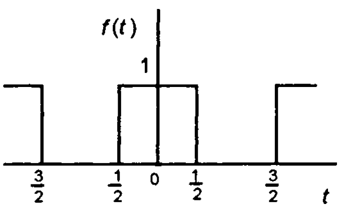


Figure 7.11 *Example*

The function is an even function and so the b coefficients are all zero. The period is 2 and so $\omega = \pi$. Thus, using equation [31]:

$$a_0 = 2\times\frac{2}{T}\int_0^{T/2} f(t)\,\mathrm{d}t = 2\left(\int_0^{1/2} 1\,\mathrm{d}t + \int_{1/2}^{1} 0\,\mathrm{d}t\right) = 2[t]_0^{1/2} = 1$$

Using equation [33]:

$$a_n = 2\times\frac{2}{T}\int_0^{T/2} f(t)\cos n\omega t\,\mathrm{d}t$$

$$= 2\left(\int_0^{1/2} 1\cos n\pi t\,\mathrm{d}t + \int_{1/2}^{1} 0\,\mathrm{d}t\right)$$

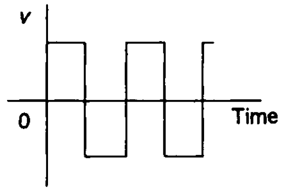
$$= 2\left[\frac{\sin n\pi t}{n\pi}\right]_0^{1/2} = \frac{2}{n\pi}\sin\tfrac{1}{2}n\pi$$

Thus the Fourier series is:

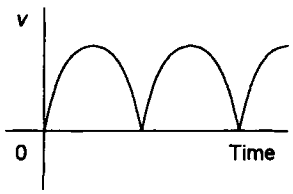
$$y = \tfrac{1}{2} + \frac{2}{\pi}\left(\cos\pi t - \tfrac{1}{3}\cos 3\pi t + \tfrac{1}{5}\cos 5\pi t + \ldots\right)$$

Half-wave symmetry

It is often possible by considering the symmetry of successive half-cycle waves within a waveform to recognise whether it will contain odd or even harmonics.



(a)



(b)

Figure 7.12 *Waveform with (a) identical positive and negative half-cycles, (b) repetition every half-cycle*

- **Half-wave inversion**
  Any complex waveform which has a negative half-cycle which is just the positive cycle inverted will contain only odd harmonics, such a form of symmetry being termed *half-wave inversion*. Thus Figure 7.12(a) shows a waveform which has negative half-cycles which are just the positive half-cycles inverted and so does not contain any even harmonics.

- **Half-wave repetition**
  Waveforms which repeat themselves after each half-cycle of the fundamental frequency will have just even harmonics, such a form of symmetry being termed *half-wave repetition*. Figure 7.12(b) show a waveform which repeats itself after each half-cycle and so has just even harmonics.

Key point

Any complex waveform which has a negative half-cycle which is just the positive cycle inverted will contain only odd harmonics, such a form of symmetry being termed *half-wave inversion*.

Waveforms which repeat themselves after each half-cycle of the fundamental frequency will have just even harmonics, such a form of symmetry being termed *half-wave repetition*.

We can see why the above statements occur by considering the conditions that are necessary for a Fourier series to give the required symmetry. Thus if we have the series describing the waveform at time t:

$$\begin{aligned} v \text{ at } t = {} & \tfrac{1}{2}a_0 + a_1 \cos 1\omega t + a_2 \cos 2\omega t + a_3 \cos 3\omega t + \dots \\ & + b_1 \sin 1\omega t + b_2 \sin 2\omega t + b_3 \sin 3\omega t + \dots \end{aligned} \qquad [37]$$

To obtain the value of the waveform after half a cycle, i.e. at time $t + \pi$, we put this value of time into equation [37]:

$$\begin{aligned} v \text{ at } (t+\pi) = {} & \tfrac{1}{2}a_0 + a_1 \cos 1\omega(t+\pi) + a_2 \cos 2\omega(t+\pi) \\ & + a_3 \cos 3\omega(t+\pi) + \dots + b_1 \sin 1\omega(t+\pi) \\ & + b_2 \sin 2\omega(t+\pi) + b_3 \sin 3\omega(t+\pi) + \dots \\ = {} & \tfrac{1}{2}a_0 - a_1 \cos 1\omega t + a_2 \cos 2\omega t - a_3 \cos 3\omega t + \dots \\ & - b_1 \sin 1\omega t + b_2 \sin 2\omega t - b_3 \sin 3\omega t + \dots \end{aligned} \qquad [38]$$

If the waveform is to have negative half-cycles which are just the positive half-cycles inverted we must have the waveform after half a cycle, i.e. at time $t + \pi$, which is $-v$ at t. Thus we must have:

$$\begin{aligned} & \tfrac{1}{2}a_0 + a_1 \cos 1\omega t + a_2 \cos 2\omega t + a_3 \cos 3\omega t + \dots \\ & \quad + b_1 \sin 1\omega t + b_2 \sin 2\omega t + b_3 \sin 3\omega t + \dots \\ & = -(\tfrac{1}{2}a_0 - a_1 \cos 1\omega t + a_2 \cos 2\omega t - a_3 \cos 3\omega t + \dots \\ & \quad - b_1 \sin 1\omega t + b_2 \sin 2\omega t - b_3 \sin 3\omega t + \dots) \end{aligned} \qquad [39]$$

This can only occur if $a_0 = 0$, $a_2 = 0$, and all even harmonics are 0.

If the waveform is to have waveforms which repeat themselves after half a cycle then we must have the waveform at time $t + \pi$ equal to v at time t. Thus we must have:

$$\begin{aligned} & \tfrac{1}{2}a_0 + a_1 \cos 1\omega t + a_2 \cos 2\omega t + a_3 \cos 3\omega t + \dots \\ & \quad + b_1 \sin 1\omega t + b_2 \sin 2\omega t + b_3 \sin 3\omega t + \dots \\ & = \tfrac{1}{2}a_0 - a_1 \cos 1\omega t + a_2 \cos 2\omega t - a_3 \cos 3\omega t + \dots \\ & \quad - b_1 \sin 1\omega t + b_2 \sin 2\omega t - b_3 \sin 3\omega t + \dots \end{aligned} \qquad [40]$$

This can only occur if $a_1 = 0$, $a_3 = 0$ and all odd harmonics are 0.

Key point

Equations [41] and [42] are for the Fourier series in the form:

$$y = A_0 + A_1 \sin(\omega t + \phi_1) + A_2 \sin(2\omega t + \phi_2) + \dots$$

Note that if we want to refer to a series in the form:

$$y = A_0 + A_1 \cos(\omega t + \phi_1) + A_2 \cos(2\omega t + \phi_2) + \dots$$

then, in order to take account of $\sin \omega t = \cos(\omega t - \pi/2)$, i.e. the phase difference of $-90°$ between the cosine and sine, equation [42] becomes:

$$\phi = \tan^{-1}\left(\frac{-b_n}{a_n}\right)$$

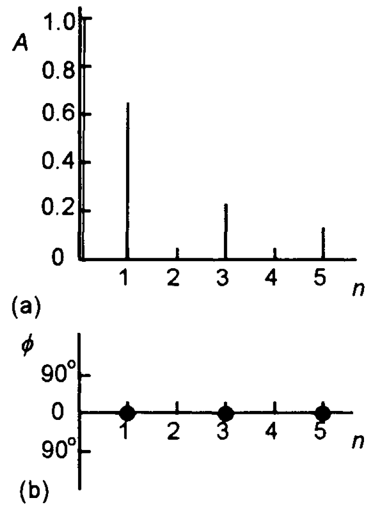


Figure 7.13 *Frequency spectrum*

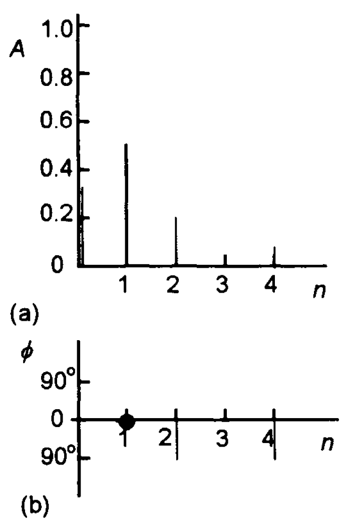


Figure 7.14 *Frequency spectrum*

7.2.4 Frequency spectrum

The *frequency spectrum* comprises an *amplitude spectrum*, which is a graph of the amplitudes of each of the constituent sinusoidal components in the Fourier series plotted against frequency, and a *phase spectrum* which is their phases. The amplitudes are given from the Fourier coefficients by equation [19]:

$$A_n = \sqrt{a_n^2 + b_n^2} \qquad [41]$$

and the phases of sinusoidal components by equation [18] as:

$$\phi_n = \tan^{-1}\left(\frac{a_n}{b_n}\right) \qquad [42]$$

Example

Determine the frequency spectrum of the rectangular waveform with $a_0 = 1$, $a_n = 0$ and $b_n = (1 - \cos n\pi)/n\pi$.

We have $b_1 = 2/\pi = 0.64$, $b_2 = 0$, $b_3 = 2/3\pi = 0.21$, $b_4 = 0$, $b_5 = 2/5\pi = 0.13$, etc. The A_0 term is 1. Using equation [41], the A_1 term is 0.64, the A_2 term 0, the A_3 term 0.21, the A_4 term 0, the A_5 term 0.13, etc. The phases, when referred to a sine wave, are 0° for all components. When referred to a cosine wave they are $-90°$. Figure 7.13(a) shows the resulting amplitude spectrum and Figure 7.13(b) the phase spectrum when referring to sinusoidal components.

Example

Determine the frequency spectrum for a half-wave rectified sinusoidal waveform if it has the Fourier series:

$$y = \frac{1}{\pi} - \frac{2}{3\pi}\cos 2\omega t - \frac{2}{15\pi}\cos 4\omega t + \dots + \frac{1}{2}\sin \omega t$$

The A_0 term is $1/\pi = 0.32$. Using equation [41], the A_1 term is 0.5, the A_2 term $2/3\pi = 0.21$, the A_3 term 0 and the A_4 term $2/15\pi = 0.04$. The phases, when referred to a sine wave are $\phi_1 = 0$ and since $-\cos \omega t = \sin(\omega t - 90°)$, $\phi_2 = -90°$ and $\phi_4 = -90°$. Figure 7.14(a) shows the amplitude spectrum and Figure 7.14(b) the phase spectrum when referring to sinusoidal components.

Maths in action

Electric circuit analysis

Often in considering electrical systems the input is not a simple d.c. or sinusoidal a.c. signal but perhaps a square wave periodic signal or a distorted sinusoidal signal or a half-wave rectified sinusoid. Such problems can be tackled by representing the waveform as a Fourier series and using the *principle of superposition*; we find the overall effect of the waveform by summing the effects due to each term in the Fourier series considered alone. Thus if we have a voltage waveform:

$$v = V_0 + V_1 \sin \omega t + V_2 \sin 2\omega t + V_3 \sin 3\omega t + \dots + V_n \sin n\omega t$$

then we can consider the effects of each element taken alone. Thus we can calculate the current due to the voltage V_0 , that due to $V_1 \sin \omega t$, that due to $V_2 \sin 2\omega t$, that due to $V_3 \sin 3\omega t$, and so on for all the terms in the series. We then add these currents to obtain the overall current due to the waveform.

Consider the application to a pure resistance R. Since $i = v/R$ and resistance R is independent of frequency, then the current due to the V_0 term is V_0/R, that due to the first harmonic term is $(V_1 \sin \omega t)/R$, that due to the second harmonic term is $(V_2 \sin 2\omega t)/R$ and so on. Thus the resulting current waveform is:

$$i = \frac{V_0}{R} + \frac{V_1}{R}\sin\omega t + \frac{V_2}{R}\sin 2\omega t + \dots + \frac{V_n}{R}\sin n\omega t$$

Because the resistance is the same for each harmonic, the amplitude of each voltage harmonic is reduced by the same factor, i.e. the resistance. The phases of each harmonic are not changed. The current waveform is thus the same shape as the voltage waveform.

Consider the application to a pure inductance L. The impedance of a pure inductance depends on the frequency, i.e. its reactance $X_L = \omega L$. Also the current lags the voltage by 90°. The impedance is 0 when the frequency is 0 and thus the current due to the V_0 term will be 0. The current due to the first harmonic will be the voltage of that harmonic divided by the impedance at that frequency and so $V_1 \sin (\omega t - 90°)/\omega L$. The current due to the second harmonic will be the voltage of that harmonic divided by the impedance at that frequency and so $V_1 \sin (2\omega t - 90°)/2\omega L$. Thus the current waveform will be:

$$i = \frac{V_1}{\omega L}\sin(\omega L - 90°) + \frac{V_2}{2\omega L}\sin(2\omega L - 90°) + \dots$$
$$+ \frac{V_n}{n\omega L}\sin(n\omega L - 90°)$$

Each of the voltage terms has its amplitude altered by a different amount; the phase, however, is changed by the same amount. The result is that the shape of the current waveform is different to that of the voltage waveform.

Consider a pure capacitor capacitance C. The impedance of a pure capacitor depends on the frequency, i.e. its reactance $X_C = 1/\omega C$, and the current leads the voltage by 90°. The impedance is 0 when the frequency is 0 and thus the current due to the V_0 term will be 0. The current due to the first harmonic will be the voltage of that harmonic divided by the impedance at that frequency and so $V_1 \sin(\omega t + 90°)/(1/\omega C)$. For the second harmonic the current will be the voltage of that harmonic divided by the impedance at that frequency and so $V_1 \sin(2\omega t + 90°)/(1/2\omega C)$. Thus the current waveform will be:

$$i = \omega C V_1 \sin(\omega t + 90°) + 2\omega C V_2 \sin(2\omega t + 90°) + \dots$$
$$+ n\omega C V_n \sin(n\omega t + 90°)$$

Each of the voltage terms has had their amplitude altered by a different amount but the phase changed by the same amount. The result of this is that the shape of the current waveform is different to that of the voltage waveform.

Example

A voltage of $2.5 + 3.2 \sin 100t + 1.6 \sin 200t$ V is applied across a resistor having a resistance of 100 Ω. Determine the current through the resistor.

The complex current will be the sum of the currents due to each of the voltage terms in the complex voltage. Since the resistance is the same at all frequencies, the complex current will be:

$i = 0.025 + 0.032 \sin 100t + 0.016 \sin 200t$ A

Thus, each of the elements has the same phase as the corresponding voltage element.

Example

A complex voltage of $2.5 + 3.2 \sin 100t + 1.6 \sin 200t$ V is applied across a pure inductor having an inductance of 100 mH. Determine the current through the inductor.

The impedance is 0 when the frequency is 0 and thus the current due to the 2.5 V term will be 0. For the second term, the reactance is $100 \times 0.100 = 10\ \Omega$ and the current lags the voltage by 90° and so the current due to this harmonic is $0.32 \sin (100t - 90°)$ A. For the third term, the reactance is $200 \times 0.100 = 20\ \Omega$ and the current lags the voltage by 90° and so the current due to this harmonic is $0.08 \sin (100t - 90°)$ A. Thus the current waveform is:

$$i = 0.32 \sin (100t - 90°) + 0.08 \sin (100t - 90°) \text{ A}$$

Example

Determine the waveform of the current occurring when a 2μF capacitor has connected across it the half-wave rectified sinusoidal voltage $v = 0.32 + 0.5 \cos 100t + 0.21 \cos 200t$ V.

There will be no current arising from the d.c. term. For the first harmonic the reactance is $1/(2 \times 10^{-6} \times 100)\ \Omega$ and so we have a current of $0.5 \times 2 \times 10^{-6} \times 100 \cos (100t + 90°)$ A. For the second harmonic the reactance is $1/(2 \times 10^{-6} \times 200)\ \Omega$ and so the current is $0.21 \times 2 \times 10^{-6} \times 200 \cos (200t + 90°)$. Thus the resulting current is:

$$i = 2 \times 10^{-6} \times 0.5 \times 100 \cos (100t + 90°)$$
$$+ 2 \times 10^{-6} \times 0.21 \times 200 \cos (200t + 90°) \text{ A}$$

Example

A voltage of $v = 100 \cos 314t + 50 \sin(5 \times 314t - \pi/6)$ V is applied to a series circuit consisting of a 10 Ω resistor, a 0.02 H inductor and a 50 μF capacitor. Determine the circuit current.

For the first harmonic, the resistance is 10 Ω, the inductive reactance is $\omega L = 314 \times 0.02 = 6.28\ \Omega$ and the capacitive reactance is $1/\omega C = 1/(314 \times 50 \times 10^{-6}) = 63.8\ \Omega$. Thus the total impedance is:

$$Z_1 = 10 + j6.28 - j63.8 = 10 - j57.52$$

$$= \sqrt{10^2 + 57.52^2} \angle \tan^{-1}\frac{-57.52}{10} = 58.4\angle(-80.1°)\ \Omega$$

Thus the current due to the first harmonic is:

$$i_1 = \frac{100\angle 0°}{58.4\angle(-80.1°)} = 1.71\angle 80.1°\ \text{A}$$

For the fifth harmonic, the resistance is 10 Ω, the inductive reactance is $5\omega L = 5 \times 314 \times 0.02 = 31.4\ \Omega$ and the capacitive reactance is $1/5\omega C = 1/(5 \times 314 \times 50 \times 10^{-6}) = 12.76\ \Omega$. Thus the total impedance is:

$$Z_5 = 10 + j31.4 - j12.76 = 10 + j18.64$$

$$= \sqrt{10^2 + 18.64^2} \angle \tan^{-1}\frac{18.64}{10} = 21.2\angle 61.8°\ \Omega$$

Thus the current due to the fifth harmonic is:

$$i_5 = \frac{50\angle(-30°)}{21.2\angle 61.8°} = 2.36\angle(-91.8°)\ \text{A}$$

Thus the current waveform is:

$$i = 1.71\cos(314t + 80.1°) + 2.36\cos(3 \times 314t - 91.8°)\ \text{A}$$

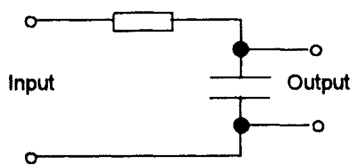


Figure 7.15 *Example*

Example

A half-wave rectified sinusoidal voltage:

$$v = 0.32 + 0.5\sin \pi t - 0.21\cos 2\pi t - 0.04\cos 4\pi t\ \text{V}$$

is applied to a circuit consisting of a 1 Ω resistor in series with a 1 F capacitor. Determine the waveform of the voltage output across the capacitor.

Figure 7.15 shows the circuit. The output is the fraction of the input voltage that is across the capacitor. Thus, using phasors and the component values given:

$$\mathbf{V}_{out} = \frac{1/jn\omega C}{(1/jn\omega C) + R}\mathbf{V}_{in} = \frac{1}{1 + jn\omega CR}\mathbf{V}_{in} = \frac{1}{1 + jn\omega}\mathbf{V}_{in}$$

where $\omega = \pi$ is the fundamental frequency. For the d.c. component, with $\omega = 0$, we have $\mathbf{V}_{out\,0} = \mathbf{V}_{in\,0} = 0.32$ V. For the first harmonic we have $\mathbf{V}_{in\,1} = -j0.5$ V and thus the output due to this term is:

$$\mathbf{V}_{out\,1} = \frac{1}{1 + j\pi}(-j0.5) = \frac{1 - j\pi}{(1 + j\pi)(1 - j\pi)}(-j0.5) = \frac{-0.5\pi - j0.}{1 + \pi^2}$$

$$= \sqrt{\frac{0.5^2\pi^2 + 0.5^2}{(1+\pi^2)^2}} \angle \tan^{-1}\frac{-0.5}{-0.5\pi} = 0.15\angle(-162.3^\circ) \text{ V}$$

For the second harmonic we have $\mathbf{V}_{\text{in }2} = -\text{j}0.21$ V and thus the output due to this term is:

$\mathbf{V}_{\text{out }2} =$

$$\frac{1}{1+\text{j}2\pi}(-\text{j}0.21) = \frac{-0.42\pi - \text{j}0.21}{1+4\pi^2} = 0.033\angle(-189^\circ)$$

For the fourth harmonic we have $\mathbf{V}_{\text{in }4} = -\text{j}0.04$ V and thus the output due to this term is:

$\mathbf{V}_{\text{out }4} =$

$$\frac{1}{1+\text{j}4\pi}(-\text{j}0.04) = \frac{-0.16\pi - \text{j}0.04}{1+16\pi^2} = 0.003\,2\angle(-184.5^\circ)$$

Thus the output is:

$$\mathbf{V}_{\text{out}} = 0.32 + 0.15\sin(\pi t - 162.3^\circ) + 0.033\cos(2\pi t - 189^\circ) + 0.003\,2\cos(4\pi t - 184.5^\circ) \text{ V}$$

Maths in action

Rectifier filter circuit

A full-wave rectifier produces a far-from-smooth output and relies on the use of a *LC* filter in order to give an output which reasonably approximates to a smooth d.c. voltage. Figure 7.16 shows the circuit. The output from the full-wave rectifier can be described by the Fourier series:

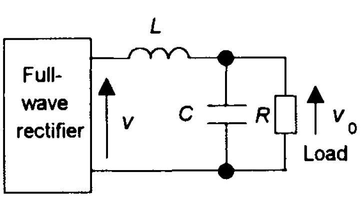


Figure 7.16 *Full-wave recifier with filter*

$$v = \frac{2V_m}{\pi}\left(1 - \tfrac{2}{3}\cos 2\omega t - \tfrac{2}{15}\cos 4\omega t - \tfrac{2}{35}\cos 6\omega t - \dots\right)$$

The first term is a constant and so represents a d.c. component. The second, and succeeding terms, represent alternating voltages which can be considered to be superimposed on the d.c. voltage.

The output voltage from filter circuit is across the resistive load. Assuming ideal components, we have a d.c. voltage of $2V_m/\pi$ across the load resistor. For the capacitor to provide effective smoothing of the output, its reactance must be low compared with the load resistance so as to divert most of the a.c. element away from the load resistor. For the a.c. elements, the circuit is effectively just a pure inductance in series with a pure capacitor. It is a voltage-divider circuit, thus for the *n*th harmonic:

$$\frac{\mathbf{V_0}}{\mathbf{V}} = \frac{-jX_C}{jX_L - jX_C}$$

and so:

$$\frac{v_0}{v} = \frac{X_C}{X_C - X_L}$$

For the 2nd harmonic we have $X_C = 1/2\omega C$ and $X_L = 2\omega L$. Thus:

$$\frac{v_0}{v} = \frac{1/2\omega C}{(1/2\omega C) - 2\omega L} = \frac{1}{1 - 4\omega^2 LC}$$

Since $4\omega^2 LC$ will be much greater than 1, the equation approximates to:

$$\frac{v_0}{v} = -\frac{1}{4\omega^2 LC}$$

For the 2nd harmonic $v = -(2V_m/\pi)(2/3)\cos 2\omega t$ and so:

$$v_0 = \frac{V_m}{3\pi\omega^2 LC}\cos 2\omega t$$

This will give a ripple on the output d.c. voltage. The size of the ripple is the peak-to-peak value of the alternating component and so is 2 times the maximum amplitude of the ripple:

$$\text{size of ripple disturbance} = \frac{2V_m}{3\pi\omega^2 LC}$$

A measure of the smoothness of the d.c. output is provided by the *ripple factor* r. This can be defined as:

$$r = \frac{\text{ripple voltage at load}}{\text{average load voltage}}$$

Thus, since we have a d.c. voltage component of $2V_m/\pi$:

$$r = \frac{1}{3\omega^2 LC}$$

As an illustration, consider the inductance needed with such a filter circuit to give a 1% ripple factor for a frequency of 50 Hz and a smoothing capacitor of 10 μF. Using these values in the above equation:

$$L = \frac{1}{3r\omega^2 C} = \frac{1}{3 \times 0.01 \times 4\pi^2 50^2 \times 10 \times 10^{-6}} = 3.38\ \text{H}$$

Problems 7.2

1 What harmonics are present in the waveform given by $v = 1.0 - 0.67 \cos 2\omega t - 0.13 \cos 4\omega t$?

2 Determine the Fourier series for the waveform shown in Figure 7.17.

3 Determine the Fourier series for the waveform in Figure 7.18.

4 Determine the Fourier series for the full-wave rectified sinusoid (Figure 7.19).

5 Determine the nature of the terms within the Fourier series for the waveforms shown in Figure 7.20. T is the periodic time for a cycle.

6 Given that the Fourier series for the waveform in Figure 7.21(a) is:

$$y = \frac{4A}{\pi}\left[\sin \omega t + \tfrac{1}{3}\sin 3\omega t + \tfrac{1}{5}\sin 5\omega t + \dots\right]$$

Determine, by considering the shift of origin, the Fourier series for the waveform shown in Figure 7.21(b).

7 What terms will be present in the Fourier series for the waveforms shown in Figure 7.22?

8 From considerations of the mean values of the waveforms in Figure 7.23, what will be the values of a_0?

9 Determine whether the following are even or odd functions:

(a) $\sin x$, (b) x, (c) $x \sin x$, (d) $x \cos 2x$, (e) $x^3 \cos 2x$

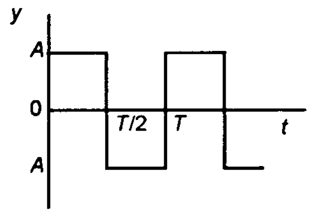


Figure 7.17 *Problem 2*

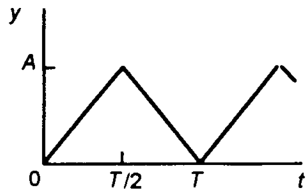


Figure 7.18 *Problem 3*

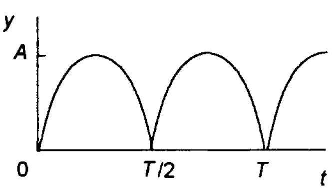


Figure 7.19 *Problem 4*

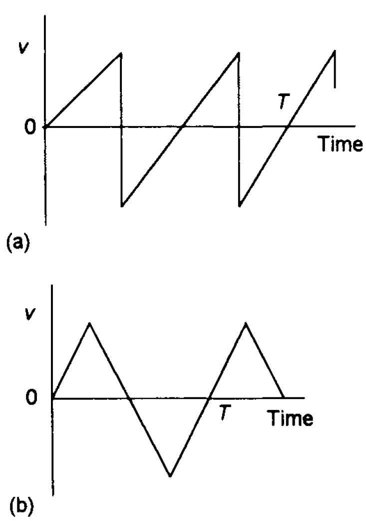


Figure 7.20 *Problem 5*

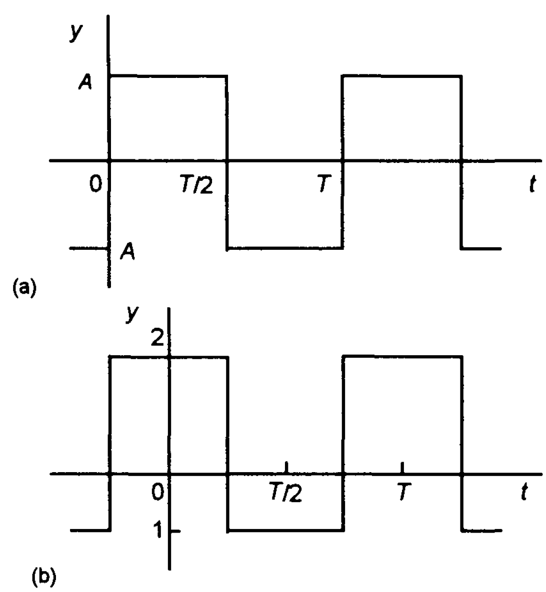


Figure 7.21 *Problem 6*

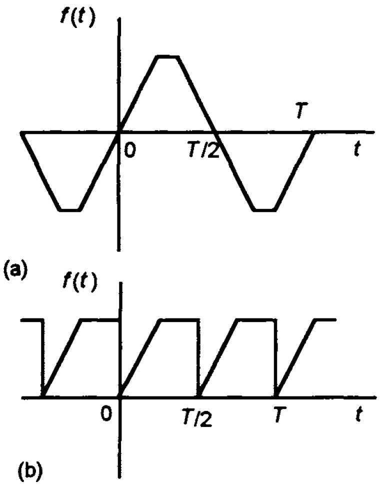


Figure 7.22 *Problem 7*

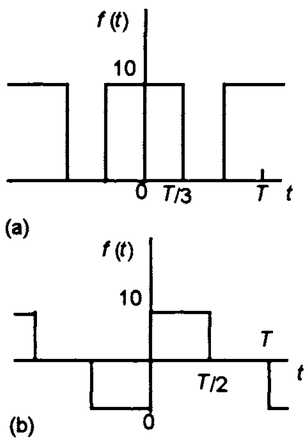


Figure 7.23 *Problem 8*

10 Determine what terms the following waveforms will contain in their Fourier series:

(a) $f(t) = 3t$ for $-\pi \le t < \pi$, period 2π,

(b) $f(t) = \cos t$ for $-\pi \le t < \pi$, period 2π,

(c) $f(t) = t^2 \cos t$ for $-\pi \le t < \pi$, period 2π

11 Determine the Fourier series for the waveform described by $f(t) = t$ for $-\pi \le t <$ o with a period of 2π.

12 Determine the amplitude and phase (referred to a sine) elements for the frequency spectrum of the waveforms giving the following Fourier series:

(a) $y = \frac{1}{2} - \frac{1}{\pi}\sin\omega t - \frac{1}{2\pi}\sin 2\omega t - \frac{1}{3\pi}\sin 3\omega t - \dots$,

(b) $a_0 = \pi/2$, $a_n = 0$ for n even and $-2/n^2\pi$ for n odd, $b_n = -(-1)^n/n$

13 Determine the waveform of the current occurring when a resistor of resistance 1 kΩ has connected across it the half-wave rectified sinusoidal voltage $v = 0.32 + 0.5 \cos 100t + 0.21 \cos 200t$ V.

14 Determine the waveform of the current when a pure inductor of inductance 10 mH has connected across it the half-wave rectified sinusoidal voltage $v = 0.32 + 0.5 \cos 100t + 0.21 \cos 200t$ V.

15 A voltage of $2.5 + 3.2 \sin 100t + 1.6 \sin 200t$ V is applied across a 10 μF capacitor. Determine the current.

8 Logic gates

Summary

In digital circuits extensive use is made of switching circuits. A switch is either on or off with these states being denoted by the digits 1 or 0. A logic circuit can be considered as a collection of switching circuits. In this chapter the basic mathematics necessary to analyse and synthesise such circuits is introduced. The mathematics involved is named after George Boole (1815–64) who first developed the modern ideas of the mathematics concerned with the manipulation of logic statements. In this chapter, Boolean algebra is approached by means of the analysis of switching circuits.

Objectives

By the end of this chapter, the reader should be able to:

- represent switching systems by logic gates;
- represent the action of such gates by truth tables;
- describe switching logic by Boolean statements;
- manipulate Boolean statements by the use of the rules of Boolean algebra.

8.1 Logic gates

Digital electronic logic gates are relatively cheap and readily available as integrated circuits. Such gates find a wide range of applications. For example, they might be used to determine when an input signal control system is to be allowed to give an output, as in an alarm system. Such logic gates are essentially just switching devices and this section considers the basics of such devices.

8.1.1 Switching circuits

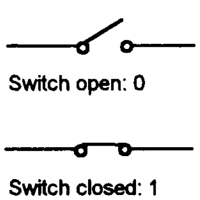


Figure 8.1 *The two states of a switch*

Consider a simple on-off switch (Figure 8.1). If we denote a closed contact by a 1 and an open contact by a 0 then the switch has just two possible states: 1 or 0.

Suppose we have two switches a and b in series. Each switch has two possible states, 0 and 1. Figure 8.2 shows the various possibilities for switches. In (a) both switches are open, in (b) a is open and b is closed, in (c) a is closed and b is open and in (d) a and b are both closed. With (a) the effect of both switches being

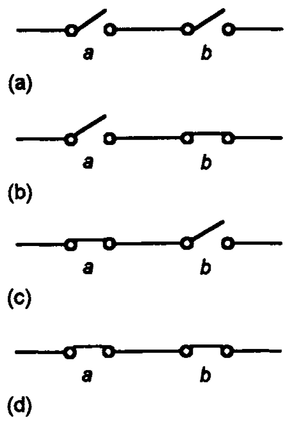


Figure 8.2 *Switches in series*

Key point

A truth table lists the outputs for each combination of inputs.

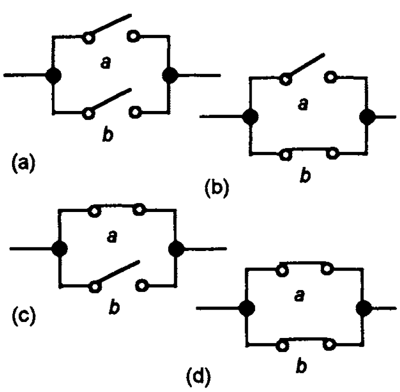


Figure 8.3 *Parallel switches*

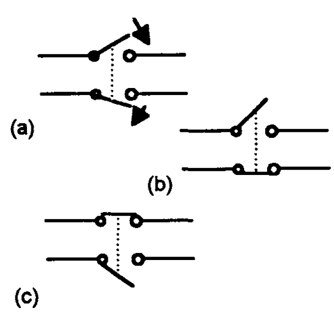


Figure 8.4 *Complement*

open is the same as would be obtained by a single open switch; (b) and (c) likewise are equivalent to a single open switch but (d) is equivalent to a single closed switch. Thus we can say that the two elements are equivalent to 0 for (a), (b) and (c) but 1 for (d). In tabular form we can represent the state of the circuit by Table 8.1:

Table 8.1 *Truth table for a* AND *b*

a	b	Output
0	0	0
0	1	0
1	0	0
1	1	1

Such a table is known as a *truth table*. If a AND b are 1 then the result is 1. Such an arrangement is known as an AND gate since both a and b have to be 1 for the output to be 1.

Consider two switches in parallel. Figure 8.3 shows the various possibilities for switches. In (a) both switches are open, in (b) a is open and b is closed, in (c) a is closed and b is open and in (d) a and b are both closed. With (a) the effect of both switches being open is the same as would be obtained by a single open switch; (b), (c) and (d) are equivalent to a single closed switch. Thus we can say that the two elements are equivalent to 0 for (a), and 1 for (b), (c) and (d). In tabular form we can represent the state of the circuit by the truth table (Table 8.2):

Table 8.2 *Truth table for a* OR *b*

a	b	Output
0	0	0
0	1	1
1	0	1
1	1	1

Such an arrangement is known as an OR gate since if a or b is 1 then the result is 1.

Another possible form of switch circuit is where two switches are connected together so that the closing of one switch results in the opening of the other. Figure 8.4(a) illustrates the switch action with (b) showing the upper switch open when the lower switch is closed and (c) the upper switch closed when the lower switch is open. The lower switch is said to give the *complement* of the upper switch. Table 8.3 is the truth table:

Table 8.3 *Truth table for* NOT

Upper switch	Lower switch
0	1
1	0

Such an arrangement constitutes a NOT switching circuit, since if one switch is 1 then the other switch is not 1.

Logic gates

With a mechanical switch we can represent the two logical states of 0 or 1 as the switch being open and closed. With electronic switches, 0 is taken to be a low voltage level and 1 a high voltage level for what is called *positive logic*, although the opposite convention (*negative logic*) can be used with 0 being represented by a high voltage level and 1 by a low voltage level. The 0 and the 1 do <u>not</u> represent actual numbers but the state of the voltage or current. The term *logic level* is often used with the voltage being said to be at logic level 0 or logic level 1.

The basic building blocks of digital electronic circuits are called *logic gates*. A logic gate is an electronic block which has one or more inputs and an output. The output can be either high or low depending on the digital levels at the input terminals. The following sections take a look at the logic gates: AND, OR, INVERT/NOT, NAND, NOR and XOR. Different sets of standard circuit symbols have been developed in Britain, Europe and the United States; an international standard (IEEE/ANSI) has, however, been developed based on squares. In this text, both the IEEE/ANSI form and the older United States form are shown.

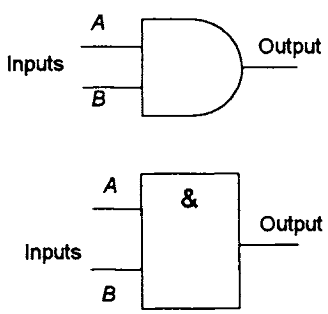


Figure 8.5 AND

- **AND** ***gate***
  The AND gate gives an output 1 when both input A <u>and</u> input B are 1. Figure 8.5 shows the symbol, the associated truth table being given in Table 8.4.

Table 8.4 AND *gate*

A	B	Output
0	0	0
0	1	0
1	0	0
1	1	1

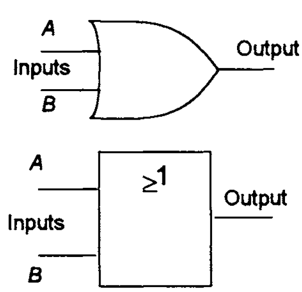


Figure 8.6 OR

- **OR** ***gate***
  The OR gate gives an output 1 when either input A <u>or</u> input B is 1. Figure 8.6 shows the symbol and Table 8.5 the truth table.

Table 8.5 OR *gate*

A	B	Output
0	0	0
0	1	1
1	0	1
1	1	1

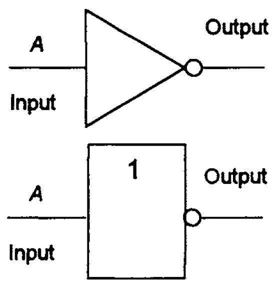


Figure 8.7 NOT

- **INVERT/NOT *gate***

The INVERT or NOT gate has a single input and gives a 1 output when the input is 0. The gate inverts the input, giving a 1 when the input is 0 and a 0 when the input is 1. Figure 8.7 shows the gate symbol and Table 8.6 gives the truth table.

Table 8.6 NOT *gate*

A	Output
0	1
1	0

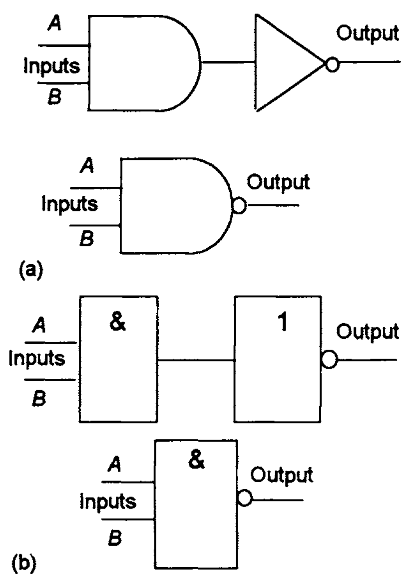


Figure 8.8 NAND

- **NAND *gate***

This gate (Figure 8.8) is logically equivalent to a NOT gate in series with an AND gate, NAND standing for NotAND. The symbol for the gate is the AND symbol followed by a small circle, the small circle being used to indicate negation. The gate has the truth table shown in Table 8.7. There is a 1 output when A and B are both not 1, i.e. are both 0.

Table 8.7 NAND *gate*

A	B	Output
0	0	1
0	1	1
1	0	1
1	1	0

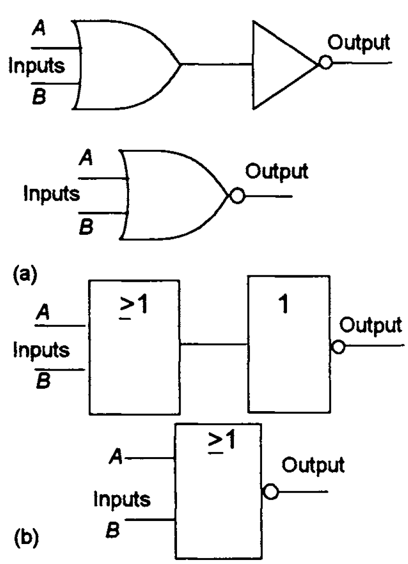


Figure 8.9 NOR

- **NOR *gate***

This gate (Figure 8.9) is logically equivalent to a NOT gate in series with an OR gate. It is represented by the OR gate symbol followed by a small circle to indicate negation. Table 8.8 gives the truth table, there being a 1 output when neither A nor B is 1.

Table 8.8 NOR *gate*

A	B	Output
0	0	1
0	1	0
1	0	0
1	1	0

- **EXCLUSIVE OR (XOR) *gate***

The OR gate gives an output 1 when either input A or input B is 1 or both A and B are 1. The EXCLUSIVE OR gate gives an output 1 when either input A or input B is 1 but not when both are. Figure 8.10 shows the gate symbol and Table 8.9 the truth table.

Table 8.9 XOR *gate*

A	*B*	Output
0	0	0
0	1	1
1	0	1
1	1	0

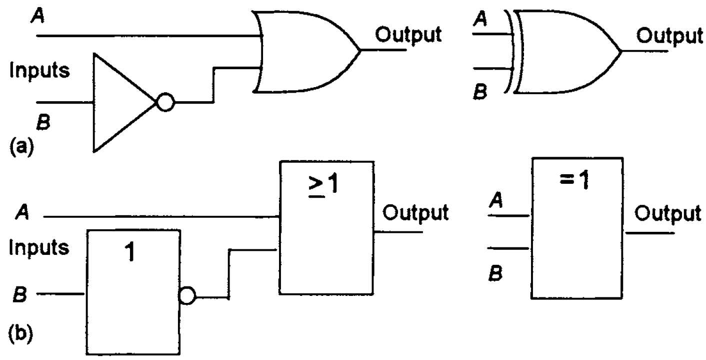


Figure 8.10 XOR

By combining gates it is possible to generate other switching operations. This is illustrated in the following example and discussed later in this chapter.

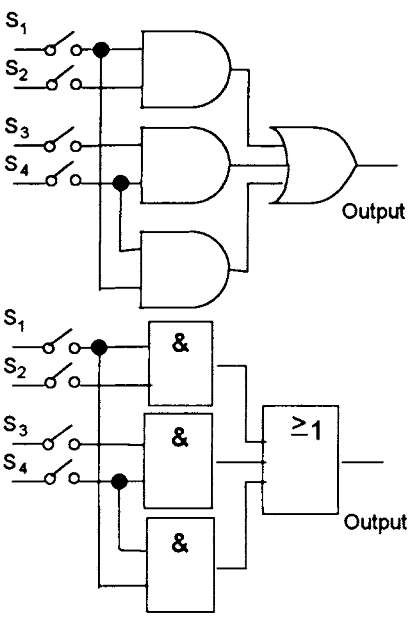


Figure 8.11 *Example*

Example

Suppose we wanted to design a switching circuit in order to operate a relay from a combination of four switches so that the relay is energised when switch 1 and switch 2 are both closed, or when switch 3 and switch 4 are both closed, or when switch 1 and switch 3 are both closed. Design a system of logic gates which would give this.

The output required is when we have (S1 and S2) or (S3 and S4) or (S1 and S3). Figure 8.11 shows how this may be realised with gates.

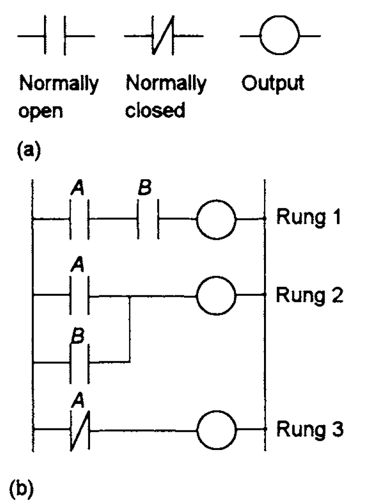


Figure 8.12 *Ladder programming*

Maths in action

Ladder programming with PLCs

Programmable logic controllers (PLCs) use a simple form of programming in order to exercise control functions. This program involves drawing each step in a program as the rung on a ladder, each rung then being taken in turn from top to bottom. Each rung can execute logic switching functions such as AND and OR. Figure 8.12(a) shows the symbols used to represent normally open switches, normally closed switches and output devices. Figure 8.12(b) shows three rungs in a ladder program. With rung 1 we have an AND gate situation in that both *A* and *B* have to be on for there to be an output. With rung 2 either *A* or *B* have to be on for there to be an output and so we have an OR gate. Rung 3 shows a NOT gate in that when *A* has an input it switches the output off.

8.2 Boolean algebra

In this section we look at how we can develop algebraic notation and rules to describe and manipulate logic gate arrangements.

Notation

For the AND operation, i.e. the series connections of switches a and b, a is considered to be *multiplied* by b. Generally $\cdot$ is used for the multiplication symbol. From truth table 8.4 we thus have the rules:

$$0 \cdot 0 = 0, \quad 0 \cdot 1 = 0, \quad 1 \cdot 0 = 0, \quad 1 \cdot 1 = 1 \qquad [1]$$

For the OR operation, i.e. the parallel connection of switches a and b, a is considered to be *added* to b. From truth table 8.5 we have the rules:

$$0 + 0 = 0, \quad 0 + 1 = 1, \quad 1 + 0 = 1, \quad 1 + 1 = 1 \qquad [2]$$

For the NOT operation, i.e. the complement with the switches, we use a bar over a symbol to indicate NOT. Thus truth table 8.6 gives the rules:

$$\overline{0} = 1, \quad \overline{1} = 0 \qquad [3]$$

Boolean algebra

The binary digits 1 and 0 are the *Boolean variables* and, together with the operations $\cdot$, + and the complement, form what is known as *Boolean algebra*. By constructing the appropriate truth tables the following laws can be derived:

Table 8.10 OR

a	a	$a + a$
0	0	0
1	1	1

Table 8.11 AND

a	a	$a \cdot a$
0	0	0
1	1	1

Table 8.12 OR

a	b	$a + b$
0	0	0
0	1	1
1	0	1
1	1	1

- ***Anything* OR*ed with itself is equal to itself***
  See Table 8.10.

  $$a + a = a \qquad [4]$$

- ***Anything* AND*ed with itself is equal to itself***
  See Table 8.11.

  $$a \cdot a = a \qquad [5]$$

- ***Input sequence for* OR *and* AND**
  It does not matter in which order we take the inputs for OR and AND gates, the output is the same. This is illustrated by Table 8.12 for OR.

  $$a + b = b + a \qquad [6]$$

  $$a \cdot b = b \cdot a \qquad [7]$$

- ***Handling bracketed terms***

As Table 8.13(a) indicates:

$$a + (b \cdot c) = (a + b) \cdot (a + c) \quad [8]$$

As Table 8.13(b) indicates.

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad [9]$$

Table 8.13(a)

a	b	c	$b \cdot c$	$a + b \cdot c$	$a + b$	$a + c$	$(a + b) \cdot (a + c)$
0	0	0	0	0	0	0	0
0	0	1	0	0	0	1	0
0	1	0	0	0	1	0	0
0	1	1	1	1	1	1	1
1	0	0	0	1	1	1	1
1	0	1	0	1	1	1	1
1	1	0	0	1	1	1	1
1	1	1	1	1	1	1	1

Table 8.13(b)

a	b	c	$b + c$	$a \cdot (b + c)$	$a \cdot b$	$a \cdot c$	$a \cdot b + a \cdot c$
0	0	0	0	0	0	0	0
0	0	1	1	0	0	0	0
0	1	0	1	0	0	0	0
0	1	1	1	0	0	0	0
1	0	0	0	0	0	0	0
1	0	1	1	1	0	1	1
1	1	0	1	1	1	0	1
1	1	1	1	1	1	1	1

Table 8.14

a	$\overline{a}$	$a + \overline{a}$
0	1	1
1	0	1

Table 8.15

a	$\overline{a}$	$a \cdot \overline{a}$
0	1	0
1	0	0

Table 8.16

a	$a + 0$	$a + 1$
0	0	1
1	1	1

- ***Complementary law***

  Anything ORed with its own negative is 1. See Table 8.14.

$$a + \bar{a} = 1 \quad [10]$$

Anything ANDed with its own negative is 0. See Table 8.15.

$$a \cdot \bar{a} = 0 \quad [11]$$

- **OR*ing with 0 or 1***

  Anything ORed with a 0 is equal to itself, anything ORed with a 1 is equal to 1. See Table 8.16.

$$a + 0 = a, \quad a + 1 = 1 \quad [12]$$

- **AND*ing with 0 or 1***

Table 8.17

a	$a \cdot 0$	$a \cdot 1$
0	0	0
1	0	1

Anything ANDed with a 0 is equal to 0, any thing ANDed with a 1 is equal to itself. See Table 8.17.

$$a \cdot 1 = a, \quad a \cdot 0 = 0 \qquad [13]$$

De Morgan laws:

- The complement of the outcome of switches a and b in parallel, i.e. an OR situation, is the same as when the complements of a and b are separately combined in series, i.e. the AND situation. Table 8.18 shows the validity of this.

$$\overline{a+b} = \overline{a} \cdot \overline{b} \qquad [14]$$

- The complement of the outcome of switches a and b in series, i.e. the AND situation, is the same as when the complements of a and b are separately considered in parallel, i.e. the OR situation. Table 8.19 shows the validity of this.

$$\overline{a \cdot b} = \overline{a} + \overline{b} \qquad [15]$$

Key points

De Morgan laws:

$\overline{a+b} = \overline{a} \cdot \overline{b}$

$\overline{a \cdot b} = \overline{a} + \overline{b}$

Table 8.18

a	b	$a+b$	$\overline{a+b}$	$\overline{a}$	$\overline{b}$	$\overline{a} \cdot \overline{b}$
0	0	0	1	1	1	1
0	1	1	0	1	0	0
1	0	1	0	0	1	0
1	1	1	0	0	0	0

Table 8.19

a	b	$a \cdot b$	$\overline{a \cdot b}$	$\overline{a}$	$\overline{b}$	$\overline{a} + \overline{b}$
0	0	0	1	1	1	1
0	1	0	1	1	0	1
1	0	0	1	0	1	1
1	1	1	0	0	0	0

Using the rules given above, complicated switching circuits can be reduced to simpler equivalent circuits.

Example

Simplify the following Boolean function:

$f = a \cdot c + (a+b) \cdot \overline{c}$.

Using equation [9]: gives $a \cdot (b + c) = a \cdot b + a \cdot c$. Since:

$(a+b)\cdot\overline{c} = a\cdot\overline{c}+b\cdot\overline{c}$

we can write:

$f = a\cdot c + a\cdot\overline{c} + b\cdot\overline{c}$

Using equation [9] for the first two terms gives:

$f = a\cdot(c+\overline{c}) + b\cdot\overline{c}$

Then using equations [7] and [10] gives:

$f = a\cdot 1+b\cdot c = a+b\cdot\overline{c}$

Example

Simplify the function: $f = a + a\cdot b\cdot c + \overline{a}\cdot\overline{c}$.

Using equation [13] we can replace a by $a \cdot 1$. The function can then be written as:

$f = a\cdot 1 + a\cdot(b\cdot c) + \overline{a}\cdot\overline{c}$

Then using equation [9]:

$f = a\cdot(1+(b\cdot c)) + a\cdot\overline{c}$

Using the second of the equations in [12] gives $1 + (b \cdot c) = 1$ and so:

$f = a\cdot 1 + \overline{a}\cdot\overline{c}$

Since $a \cdot 1 = a$ (equation [10]), and applying equation [8]:

$f = a + \overline{a}\cdot\overline{c} = (a+\overline{a})\cdot(a+\overline{c})$

But: $a+\overline{a} = 1$

and so, using equation [13]:

$f = a+\overline{c}$

Example

The operation of an output relay controlled by a PLC program is given by the Boolean expression:

$$\begin{aligned}Y007 = {} & (X001 \cdot X002 \cdot M000 \cdot M002)\\ & + (M002 \cdot X001 \cdot \overline{X004} \cdot M000)\\ & + (X003 \cdot M003 \cdot M002 \cdot X002)\\ & + (M003 \cdot X003 \cdot M002 \cdot X006)\end{aligned}$$

(a) Represent this expression as rungs in a PLC ladder program, with a rung for each part of the expression.
(b) Simplify the ladder program and hence write another Boolean expression which describes the simplified program.

(a) Each bracketed term can be represented by a rung in a ladder program and so give the program shown in Figure 8.13(a).
(b) Figure 8.13(b) shows how we can simplify the ladder program and still give the same outcome. Such a program can then be described by the Boolean expression:

$$\begin{aligned}Y007 = {} & ((X002 + X004) \cdot X001 \cdot M000)\\ & + ((X002 + X006) \cdot X003 \cdot M003)) \cdot M002\end{aligned}$$

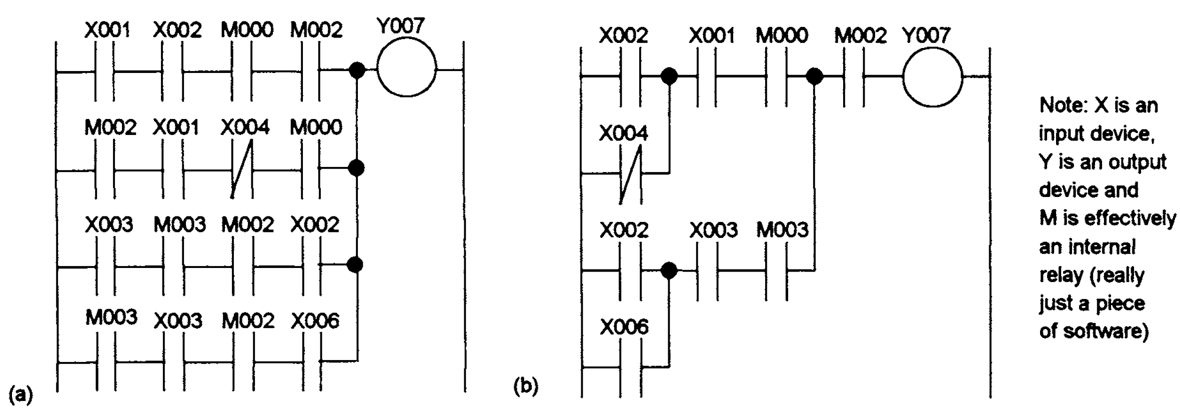


Figure 8.13 *Example*

Problems 8.2

1 Complete the following:

(a) $1 + 0 = ?$, (b) $1 \cdot 1 = ?$, (c) $\overline{1} = ?$

2 Simplify the following Boolean functions:

(a) $a(\overline{a} + a \cdot b)$, (b) $a + b + c + \overline{a} \cdot b$, (c) $(a + b) \cdot (a + b)$,

(d) $a \cdot \overline{b} \cdot c + a \cdot \overline{b} \cdot \overline{c}$, (e) $a + \overline{a} \cdot \overline{b}$

8.3 Logic gate systems

The operations ·, + and the complement can be used to write the Boolean functions for complex switching circuits, the states of such circuits being determined by developing the truth table to indicate all the various switching possibilities. Boolean algebra might then be used to simplify the switching circuits.

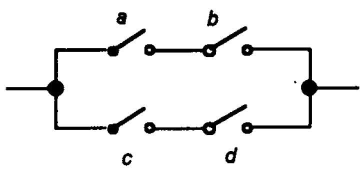


Figure 8.14 *Example*

Example

Write, for the circuit shown in Figure 8.14, (a) the truth table and (b) the Boolean function to describe that truth table.

(a) a and b are in series, and in parallel with the series arrangement of c and d. The result of using the switches is that only when either a and b are closed or c and d are closed will there be an output. Table 8.20 shows the truth table.
(b) The Boolean function for the two switches a and b in series is $a \cdot b$, the AND function, and thus, since the function for two items in parallel is OR, the function for the circuit as a whole is:

$$a \cdot b + c \cdot d$$

Table 8.20 *Example*

a	b	c	d	Result
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	1
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	1
1	1	0	0	1
1	1	0	1	1
1	1	1	0	1
1	1	1	1	1

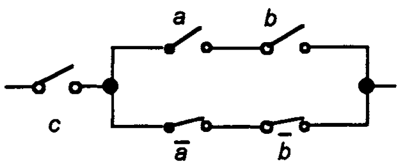


Figure 8.15 *Example*

Example

Derive the Boolean function for the switching circuit shown in Figure 8.15.

In the upper parallel arm of the circuit, the switches *a* and *b* are in series and so have a Boolean expression of $a \cdot b$. In the lower arm the complements of *a* and *b* are in series. Thus the Boolean expression for that part of the circuit is $\overline{a} \cdot \overline{b}$.

Because the two arms are in parallel the expression for the parallel part of the circuit is $a \cdot b + \overline{a} \cdot \overline{b}$.

In series with this is switch *c*. Thus the Boolean function for the circuit is: $c \cdot (a \cdot b + \overline{a} \cdot \overline{b})$.

Combining gates

By combining logic gates it is possible to represent other Boolean functions and use of Boolean algebra can often be used to simplify the arrangement.

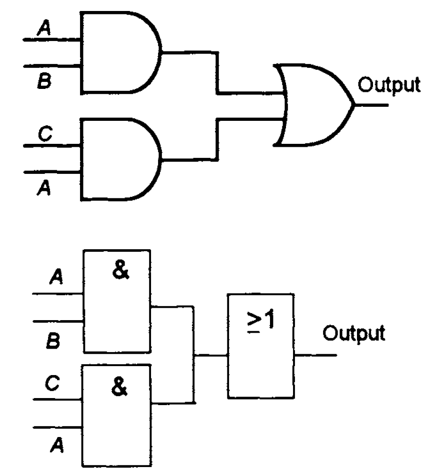


Figure 8.16 *Example*

Example

Determine the Boolean function describing the relation between the output from the logic circuit shown in Figure 8.16. Hence, consider how the circuit could be simplified.

This might be a circuit used with a car warning buzzer so that it sounds when the key is in the ignition (*A*) and a car door is opened (*B*) or the headlights are on (*C*) and a car door is opened (*A*). We have two AND gates and an OR gate. The output from the top AND gate is $A \cdot B$, and from the lower AND gate $C \cdot A$. These outputs are the inputs to the OR gate and thus the output is

output = $A \cdot B + C \cdot A$

The circuit can be simplified by considering the Boolean algebra. Using equation [9] the Boolean function can be written as:

$A \cdot B + C \cdot A = A \cdot (B + C)$

We now have *A* and *B* or *C*. This function now describes a logic circuit with just two gates, an OR gate and an AND gate. Figure 8.17 shows the circuit.

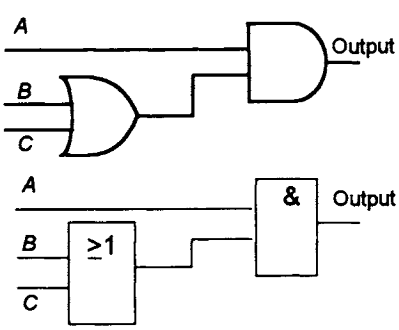


Figure 8.17 *Example*

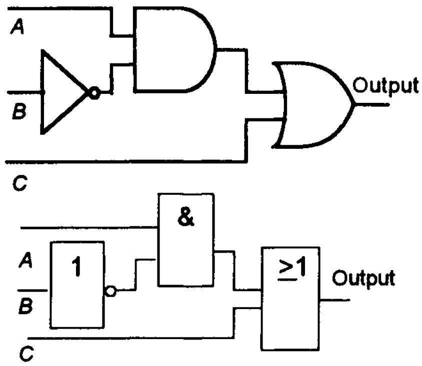


Figure 8.18 *Example*

> ### Example
>
> Devise a logic gate system to generate the Boolean function $A \cdot \overline{B} + C$.
>
> $A \cdot B$ requires an AND gate, but as the B input has to be inverted we precede the input from B to the AND gate by a NOT gate. We then require an OR gate for the output from the AND gate and C. Figure 8.18 shows the gate system.

Boolean function generation from truth tables

Often the requirements for a system are specified in terms of a truth table and the problem then becomes one of determining how a logic gate system can be devised, using the minimum number of gates, to give that truth table. The forms to which most are minimised are an AND gate driving a single OR gate or vice versa.

- ***Sum of products***

  Two AND gates driving a single OR gate (Figure 8.19(a)) give, what is termed, the sum of products form:

  $A \cdot B + A \cdot C$

- ***Product of sums***

  Two OR gates driving a single AND gate (Figure 8.19(b)) give, what is termed the product of sums form:

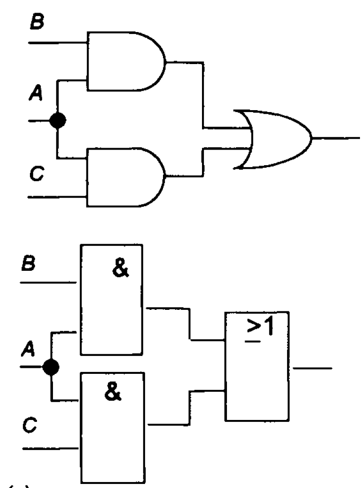
  $(A + B) \cdot (A + C)$

The usual procedure to find the minimum logic gate system is thus to find the sum of products or the product of sums form that fits the data. Generally the sum of products form is used. The procedure used is:

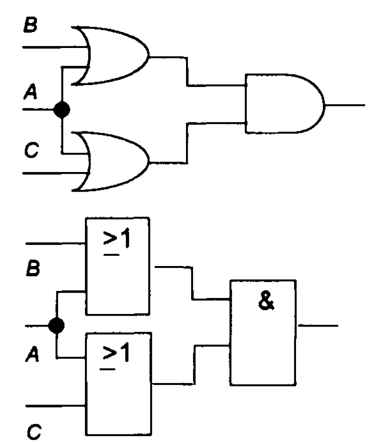
1. Consider each row of the truth table in turn that generates a 1 output and find the product that would fit a row. Only a row of a truth table that has an output of 1 need be considered, since the rows with 0 output do not contribute to the final expression. For example, suppose we have a row in a truth table of: $A = 1$, $B = 0$ and output = 1. When A is 1 and B is not 1 then the output is 1, thus the product which fits this is:

   $Q = A \cdot \overline{B}$

2. The overall result is the sum of all the products for the rows giving 1 output.



(a)



(b)

Figure 8.19 *(a) Sum of products, (b) product of sums*

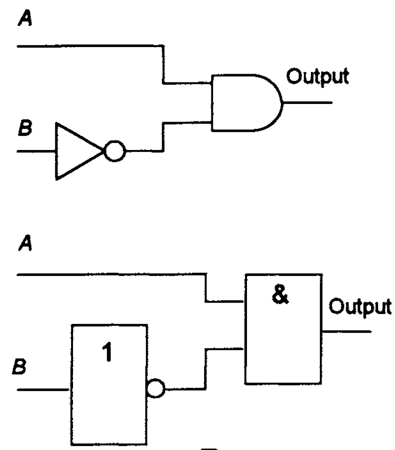


Figure 8.20 $A \cdot \overline{B}$

Example

Determine a logic gate system to give the following truth table.

A	B	Output	Products
0	0	0	
0	1	0	
1	0	1	$A \cdot \overline{B}$
1	1	0	

We only need consider the third row, thus the result is:

$$Q = A \cdot \overline{B}$$

The logic gate system that will give this truth table is thus that shown in Figure 8.20.

Example

Determine a logic gate system which will give the following truth table.

A	B	C	Output	Products
0	0	0	1	$\overline{A} \cdot \overline{B} \cdot \overline{C}$
0	0	1	0	
0	1	0	1	$\overline{A} \cdot B \cdot \overline{C}$
0	1	1	0	
1	0	0	0	
1	0	1	0	
1	1	0	0	
1	1	1	0	

There are two rows for which we need to find a product. Thus the sum of products which fits this table is:

$$Q = \overline{A} \cdot \overline{B} \cdot \overline{C} + \overline{A} \cdot B \cdot \overline{C}$$

This can be simplified to give:

$$Q = \overline{A} \cdot \overline{C} \cdot (\overline{B} + B) = \overline{A} \cdot \overline{C}$$

The truth table can thus be generated by just a NAND gate.

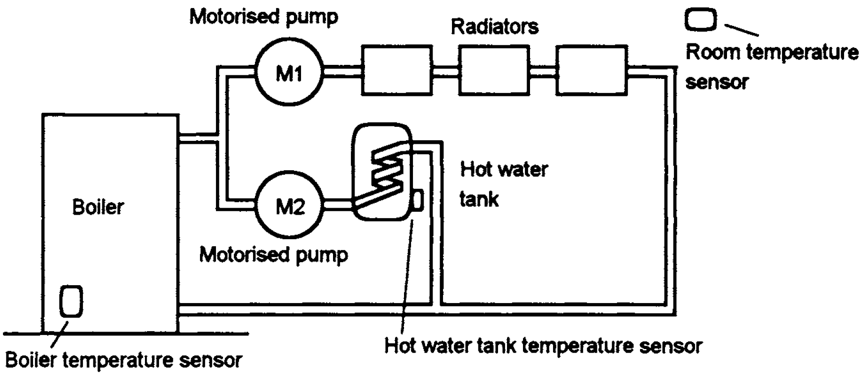


Figure 8.21 *Central heating system*

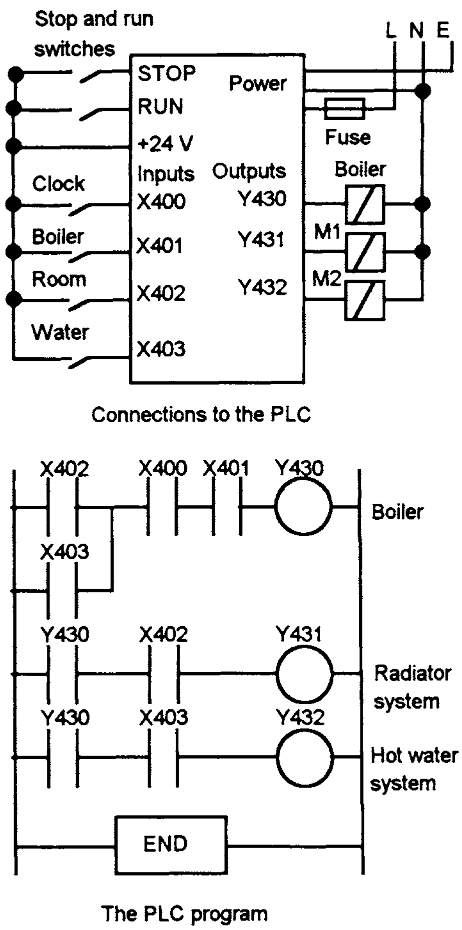


Figure 8.22 *Central heating system*

Maths in action

A PLC and a central heating system

Consider a domestic central heating system (Figure 8.21) and its control by a PLC (see earlier Maths in Action in this chapter). The central heating boiler is to be thermostatically controlled and supply hot water to the radiator system in the house and also to the hot water tank to provide hot water from the taps in the house. Pump motors have to be switched on to direct the hot water from the boiler to either, or both, of the radiator and hot water systems according to whether the temperature sensors for the room temperature and the hot water tank indicate that the radiators or tank need heating. The entire system is to be controlled by a clock so that it only operates for certain hours of the day. Figure 8.22 shows how a PLC might be used and its ladder program.

The boiler, output Y430, is switched on if X400 and X401 and either X402 or X403 are switched on. This means if the clock switched is on, the boiler temperature sensor gives an on input, and either the room temperature sensor or the water temperature sensors give on inputs. The motorised valve M1, output Y431, is switched on if the boiler, Y430, is on and if the room temperature sensor X402 gives an on input. The motorised valve M2, output Y432, is switched on if the boiler, Y430, is on and if the water temperature sensor gives an on input.

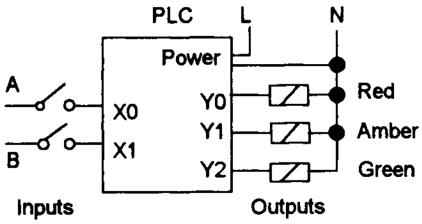


Figure 8.23 *Example*

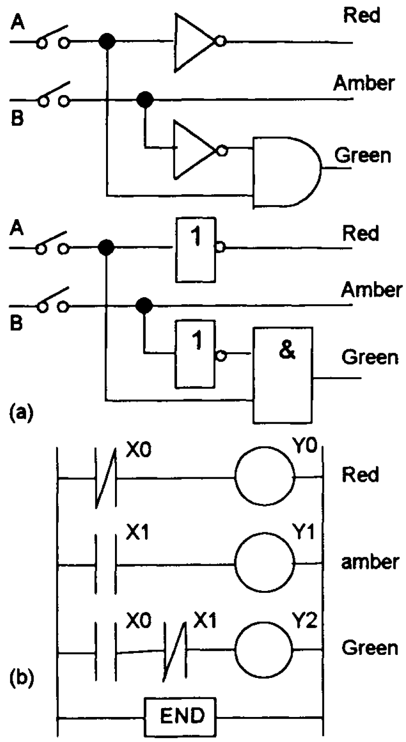


Figure 8.24 *Example*

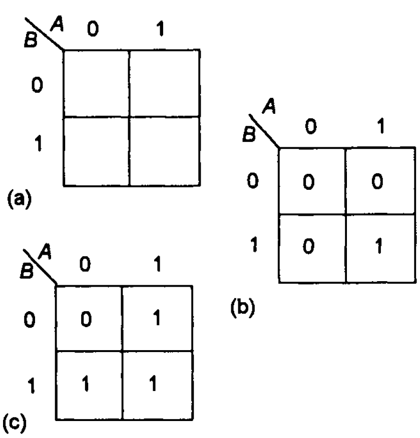


Figure 8.25 *Karnaugh map with a four-cell box*

Example

Design a PLC ladder program that will control a simple red–amber–green traffic light sequence for two inputs X0 and X1 to give the outputs Y0, Y1 and Y2 (Figure 8.23) shown in the following table:

Inputs		Outputs		
X0	X1	YO	Y1	Y2
0	0	1	0	0
0	1	1	1	0
1	0	0	0	1
1	1	0	1	0

Note: logic 0 defines an open switch or a light turned OFF, logic 1 defines a closed switch and a light turned ON.

Figure 8.24(a) shows how we can represent the above truth table by a logic gate system and Figure 8.24(b) by rungs in a ladder program.

When there is no input to X0 then the red light is ON. Thus, when the input to X0 is 0 then, as the switch is normally closed, the output Y1 is ON; when the input is 1, to open the switch, then the output is OFF.

Karnaugh maps

The *Karnaugh map* is a graphical way of representing a truth table and a method by which simplified Boolean expressions can be obtained from sums of products. The Karnaugh map is drawn as a rectangular array of cells with each cell corresponding to the output for a particular combination of inputs, i.e. a particular product value. Thus, Figure 8.25(a) shows the four-cell box corresponding to two input variables A and B, this giving four product terms. We then insert the function for each input combination, Figure 8.25(b) showing this for an AND gate and Figure 8.25(c) for an OR gate. Figure 8.26 shows how we can represent such maps with input labels A and B for 1 entries and not A and B for 0 entries.

Karnaugh maps not only pictorially represent truth tables but also can be used for minimisation. Suppose we have the following truth table:

A	B	Output	Products
0	0	0	$\overline{A} \cdot \overline{B}$
0	1	0	$\overline{A} \cdot B$
1	0	1	$A \cdot \overline{B}$
1	1	0	$A \cdot B$

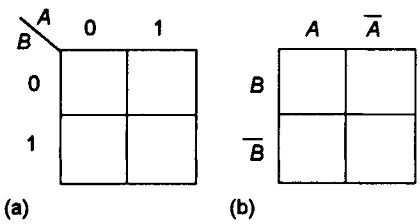


Figure 8.26 *Karnaugh map with a four-cell box*

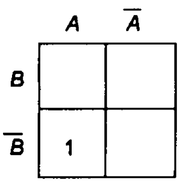


Figure 8.27 *Karnaugh map*

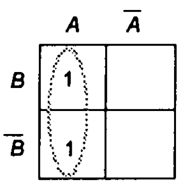


Figure 8.28 *Karnaugh map*

Key point

Simplification using Karnaugh maps

1. Construct the Karnaugh map and place 1s in those squares which correspond to the 1s in the truth table.
2. Examine the map for adjacent 1s and loop them.
3. Form the OR sum of all those terms generated by each loop.

Figure 8.27 shows the Karnaugh map for this truth table with just the 1 output shown. On the map this entry is just the cell with the coordinates $A = 1$, $B = 0$ and so gives the indicated product. The Karnaugh map enables the minimisation to be spotted visually.

As a further example, consider the following truth table:

A	B	Output	Products
0	0	0	$\overline{A} \cdot \overline{B}$
0	1	0	$\overline{A} \cdot B$
1	0	1	$A \cdot \overline{B}$
1	1	1	$A \cdot B$

Figure 8.28 shows the Karnaugh map for this truth table with just the 1 output shown. This has an output given by:

$$A \cdot \overline{B} + A \cdot B$$

This can be simplified to:

$$A \cdot \overline{B} + A \cdot B = A \cdot (\overline{B} + B) = A$$

Thus we have a rule for the map that: *when two cells containing a 1 have a common edge then we can simplify them to just the common variable, the variable that appears in the complemented and uncomplemented form is eliminated.* To help with such simplifications, we draw loops round 1s in adjacent cells. Note that in looping, adjacent cells can be considered to be those in the left- and right-hand columns. Think of the map as though it is wrapped round a vertical cylinder and the left- and right-hand edges of the map are joined together. There are other rules we can develop, namely: *looping a quad of adjacent 1s eliminates the two variables that appear in complemented and uncomplemented form* and *looping an octet of adjacent 1s eliminates the three variables that appear in both complemented and uncomplemented form.*

Figure 8.29(a) shows how we can draw a Karnaugh map for three inputs and Figure 8.29(b) for four inputs. Note that the cells are labelled so horizontally adjacent cells differ by just one variable, likewise adjacent vertical cells.

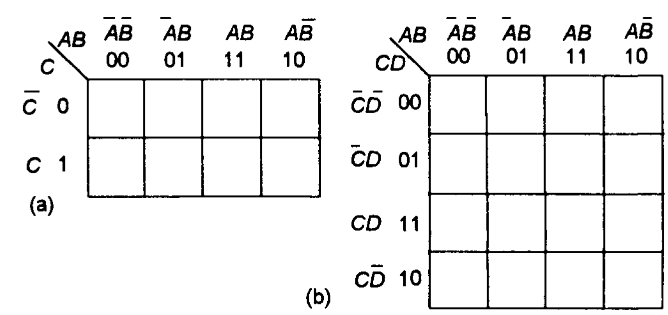


Figure 8.29 *Karnaugh maps: (a) three-input, (b) four-input*

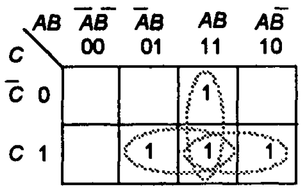


Figure 8.30 *Karnaugh map*

> ### Example
> Determine the simplified Boolean expression for the Karnaugh map shown in Figure 8.30.
>
> We have three loops and so the outcome is:
>
> $A\cdot B+B\cdot C+A\cdot C$

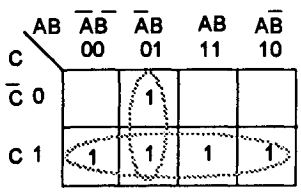


Figure 8.31 *Karnaugh map*

> ### Example
> Determine the simplified Boolean expression for the Karnaugh map shown in Figure 8.31.
>
> We have a doublet loop and a quad loop and so the outcome is:
>
> $\overline{A}\cdot B+C$

Problems 8.3

1 State a Boolean function that can be used to represent each of the switching circuits shown in Figure 8.32.

2 Give the truth tables for the switching circuits represented by the Boolean functions:

(a) $(a+\overline{b})+(a+\overline{c})$, (b) $\overline{a}\cdot(a\cdot b+\overline{b})\cdot\overline{b}$

3 Determine the Boolean functions that could generate the outputs in Figure 8.33.

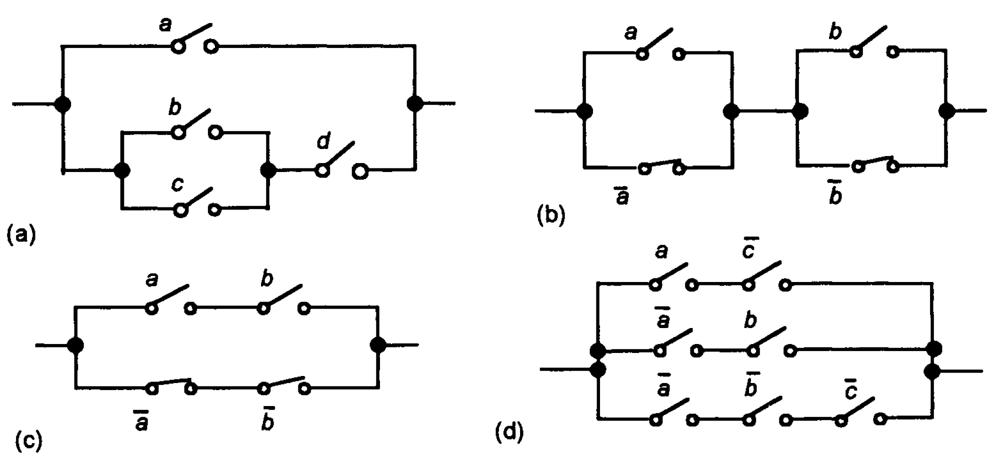


Figure 8.32 *Problem 1*

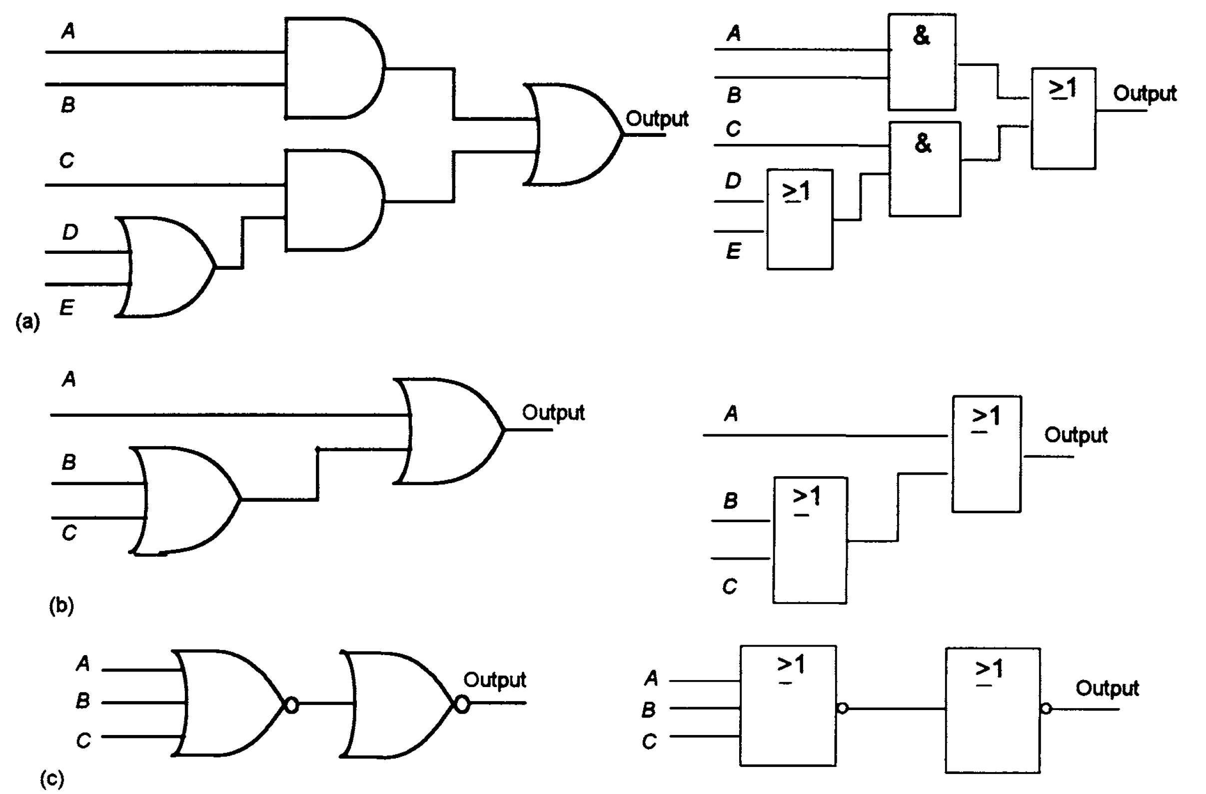


Figure 8.33 *Problem 3*

Table 8.21(a)

a	*b*	*c*	*Function*
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	1
1	1	1	0

Table 8.21(b)

a	*b*	*c*	*Function*
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

4 Give the truth table for the switching circuit corresponding to the Boolean function:

$$(a\cdot\overline{b})+(\overline{a}\cdot b)$$

5 Draw switching circuits to represent the following Boolean functions:

(a) $a\cdot(a+b)$, (b) $a\cdot(a\cdot b+c)$, (c) $\bar{a}\cdot(a+\overline{b}\cdot(a+c))$

6 Determine the Boolean equations describing the logic circuits in Figure 8.34, then simplify the equations and hence obtain simplified logic circuits.

7 Draw switching circuits to represent the Boolean functions :

(a) $a\cdot b$, (b) $a\cdot b+b$, (c) $c\cdot(a\cdot b+a\cdot\overline{b})$,
(d)$a\cdot(a\cdot b\cdot\overline{c}+a\cdot(\overline{b}+c))$.

8 Derive the Boolean functions for the truth tables in Table 8.21(a) and (b).

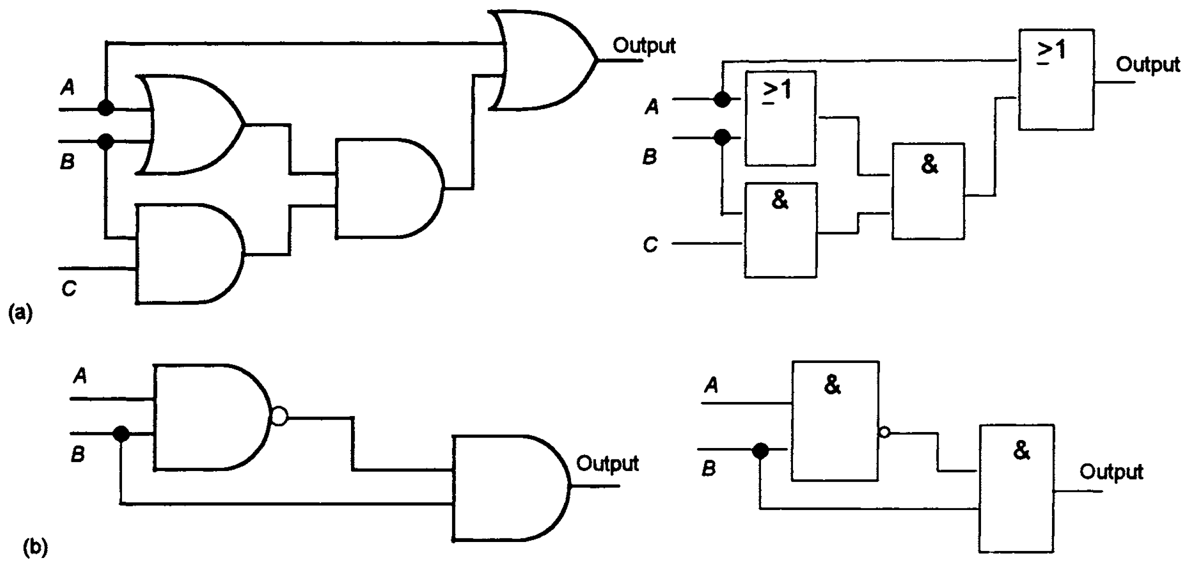


Figure 8.34 *Problem 6*

9 Determine the Boolean equations describing the logic circuits in Figure 8.35, then simplify the equations and hence obtain simplified logic circuits.

10 For the Karnaugh maps in Figure 8.36, produce the simplified Boolean expression.

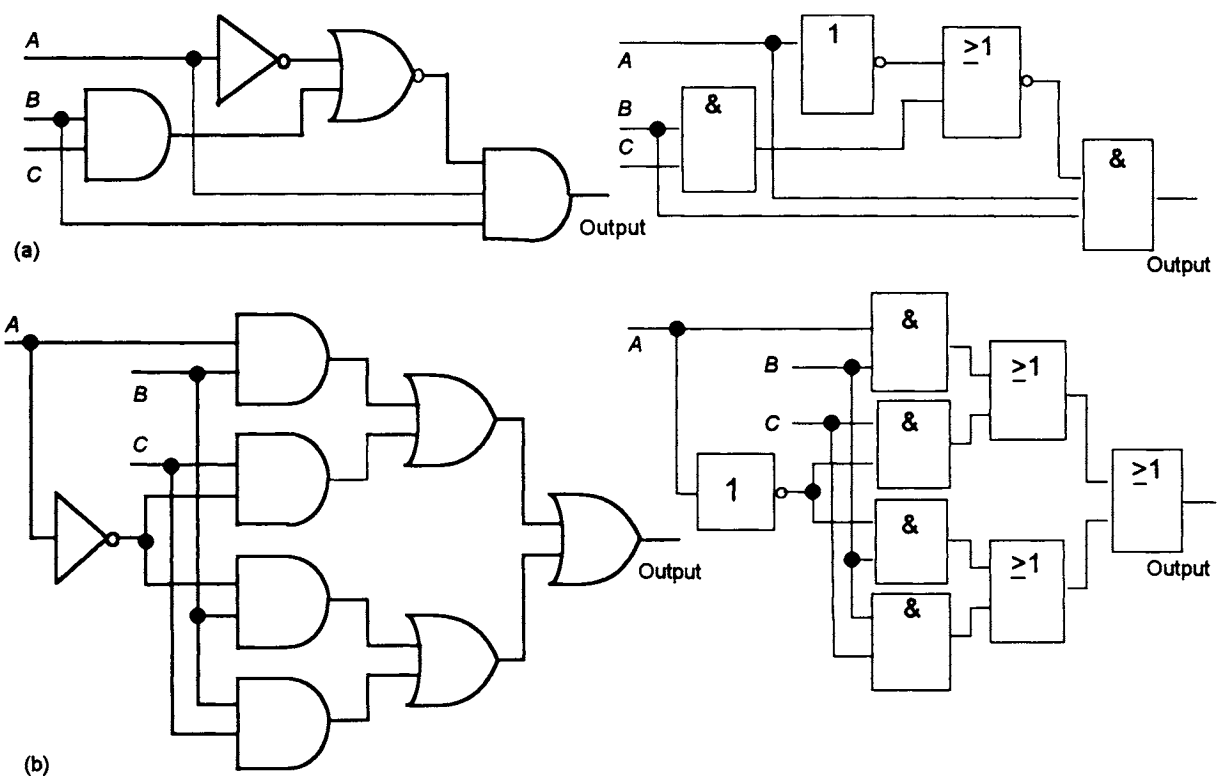


Figure 8.35 *Problem 8*

(a)

$B \backslash A$	0	1
0	1	
1		1

(b)

$C \backslash AB$	$\bar{A}\bar{B}$	$\bar{A}B$	AB	$A\bar{B}$
$\bar{C}$		1	1	
C				

(c)

$C \backslash AB$	$\bar{A}\bar{B}$	$\bar{A}B$	AB	$A\bar{B}$
$\bar{C}$	1	1		
C		1	1	1

(d)

$C \backslash AB$	$\bar{A}\bar{B}$	$\bar{A}B$	AB	$A\bar{B}$
$\bar{C}$	1			1
C				

(e)

$C \backslash AB$	$\bar{A}\bar{B}$	$\bar{A}B$	AB	$A\bar{B}$
$\bar{C}$				
C	1	1	1	1

(f)

$CD \backslash AB$	$\bar{A}\bar{B}$	$\bar{A}B$	AB	$A\bar{B}$
$\bar{C}\bar{D}$				
$\bar{C}D$				
CD	1	1	1	1
$C\bar{D}$				

(g)

$CD \backslash AB$	$\bar{A}\bar{B}$	$\bar{A}B$	AB	$A\bar{B}$
$\bar{C}\bar{D}$	1		1	1
$\bar{C}D$	1		1	1
CD			1	
$C\bar{D}$			1	

Figure 8.36 *Problem 10*

9 Probability and statistics

Summary

In any discussion of system reliability or quality control, the concept of probability plays a vital part. It is also a vital issue in the consideration of statistics when errors have to be considered in experimental measurements, all measurements being subject to some degree of uncertainty. For example, in the control of manufactured items (statistical process control SPC) control is exercised of measured variables against go/no-go criteria, the attribute, to avoid incurring scrap or reworking costs. This chapter is an introductory consideration of the principles of probability and statistics allied to such engineering issues.

Objectives

By the end of this chapter, the reader should be able to:

- understand the concept of probability;
- use probability principles in the consideration of quality control and system reliability;
- plot experimentally obtained data to show its distribution;
- use the idea of probability distributions and be able to interpret them;
- determine measures of the location and spread of distributions;
- use measures obtained from the Binomial, Poisson and Normal distributions;
- determine the errors in results obtained from experimental measurements.

9.1 Probability

What is the chance an engineering system will fail? What is the chance that a product emerging from a production line is manufactured to the required engineering tolerances, thus avoiding reworking or scrap? What is the chance that if you make a measurement in some experiment that it will be the true value of that quantity? Within what range of experimental error might you expect a measurement to be the true value? These, and many other questions in engineering and science, involve a consideration of chances of events occurring. The term *probability* is more often used in mathematics than chance and has the same meaning in the above questions. This section is about probability, its definition and determination in a number of situations.

9.1.1 Basic definitions

If you flip a coin into the air, what is the chance that it will land heads uppermost? We can try such an experiment and determine the outcomes. The result of a large number of trials leads to the result that about half the time it lands heads uppermost and half the time tails uppermost. If n is the number of trials then we can define probability P as:

$$P = \lim_{n \to \infty} \frac{\text{number of times an event occurs}}{n} \qquad [1]$$

This view of probability is *the relative frequency in the long run with which an event occurs*. In the case of the coin this leads to a probability of ½ = 0.5. If an event occurs all the time then the probability is 1. If it never occurs the probability is 0.

The result of flipping the coin might seem obvious since there are just two ways a coin can land and just one of the ways leads to heads uppermost. If there is no reason to expect one way is more likely than the other then we can define probability P as *the degree of uncertainty about the way an event can occur* and as:

$$P = \frac{\text{number of ways a particular event can occur}}{\text{total number of ways events can occur}} \qquad [2]$$

In the case of the coin, this also gives a probability of 0.5. If every possible way events can occur is the required way, then the probability is 1. If none of the possible ways are the event required, then the probability is 0.

Consider a die-tossing experiment. A die can land in six equally likely ways, with uppermost 1, 2, 3, 4, 5, or 6. Of the six possible ways the die could land, only one way is with 6 uppermost. Thus using definition [2], the probability of obtaining a 6 is 1/6. The probability of *not* obtaining a 6 is 5/6 since there are 5 ways out of the 6 possible ways we can obtain an outcome which is not a 6.

Another way the term probability is used is as *degree of belief*. Thus we might consider the probability of a particular horse winning a race as being 1 in 5 or 0.2. The probability in this case is highly subjective.

Key points

Probability can be defined as the relative frequency in the long run with which an event occurs or as the fraction of the total number of ways with which an event can occur.

Example

In the testing of products on a production line, for every 100 tested 5 were found to contain faults. What is the probability that in selecting one item from 100 on the production line that it will be faulty?

There are 100 ways the item can be selected and 5 of the ways give faulty items. Thus, using equation [2], the probability is 5/100 = 0.05.

Maths in action

The term *reliability* in relation to engineering systems is defined as the probability that the system will operate to an agreed level of performance for a specified period, subject to specified environmental conditions. Thus, for example, the reliability of an instrument might be specified as being 0.8 over a 1000 hour period with the ambient temperature at 20°C ± 10°C and no vibration.

Key points

An event which has a certainty has a probability of 1. Thus such an event will always occur every time. An event which has a probability of 0 will never occur.

Key points

Mutually exclusive events are ones for which each outcome is such that one outcome excludes the occurrence of the other.

Addition rule: If an event can happen in a number of different and mutually exclusive ways, the probability of its happening is the sum of the separate probabilities that each event happens.

Probability of events

If an event can occur in two possible ways, e.g. a piece of equipment can be either operating satisfactorily or have failed, then if the probability of one way is P_1 and the probability of the other way is P_2, we must have:

$$P_1 + P_2 = 1 \qquad [3]$$

The probability of either event 1 or event 2 occurring equals 1, i.e. a certainty, and is the sum of the probability of event 1 occurring, i.e. P_1, added to the probability of event 2, i.e. P_2, occurring.

> *A probability of 1 for an event means that the probability of it occurring is a certainty.*

Suppose with the die-tossing experiment we were looking for the probability that the outcome would be an even number. Of the six possible outcomes of the experiment, three ways give the required outcome. Thus, using definition [2], the probability of obtaining an even number is 3/6 = 0.5. This is the sum of the probabilities of 2 occurring, 4 occurring and 6 occurring, i.e. 1/6 + 1/6 + 1/6. The 2, the 4 and the 6 are mutually exclusive events in that if the 2 occurs then 4 or 6 cannot also be occurring. Thus:

> *If A and B are mutually exclusive, the probability of A or B occurring is the sum of the probabilities of A occurring and of B occurring.*

Example

The probability that a circuit will malfunction is 0.01. What is the probability that it will function?

The probability that it will function and the probability that it will not function must together be 1. Hence the probability that it will function is 0.99.

Example

A company manufactures two products *A* and *B*. Market research over a month showed 30% of enquiries by potential customers resulting in product *A* alone being bought, 50% buying product *B* alone, 10% buying both *A* and *B* and 10% buying neither. Determine the probability that an enquiry will result in (a) product *A* alone being bought, (b) product *A* being bought, (c) both product *A* and product *B* being bought, (d) product *B* not being bought.

(a) 30% buy product *A* alone so the probability is 0.30.
(b) 30% buy product *A* alone and 10% buy *A* in conjunction with *B*. Thus the probability of *A* being bought is 0.30 + 0.10 = 0.40.
(c) 10% buy products *A* and *B* so the probability is 0.10.
(d) 50% buy product *B* alone and 10% buy *B* in conjunction with *A*. Thus the probability of buying *B* is 0.50 + 0.10 = 0.60. The probability of not buying *B* is thus 1 – 0.60 = 0.40.

9.1.2 Ways events can occur

Suppose we flip two coins. What is the probability that we will end up with both showing heads uppermost? The ways in which the coins can land are:

HH HT TH TT

There are four possible results with just one of the ways giving HH. Thus the probability of obtaining HH is ¼ = 0.25.

There were two possible outcomes from the experiment of tossing the first coin and two possible outcomes from the experiment of tossing the second coin. For each of the outcomes from the first experiment there were two outcomes from the second experiment. Thus for the two experiments the number of possible outcomes is 2 × 2 = 4. This is an example of, what is termed, the *multiplication rule*.

Tree diagrams can be used to visualise the outcomes in such situations, Figure 9.1 showing this for the two experiments of tossing coins.

Key point

Multiplication rule: If one experiment has n_1 possible outcomes and a second experiment n_2 possible outcomes then the compound experiment of the first experiment followed by the second has $n_1 \times n_2$ possible outcomes.

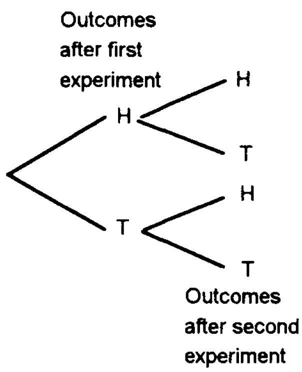


Figure 9.1 *Tree diagram*

Example

A company is deciding to build two new factories, one of them to be in the north and one in the south. There are four potential sites in the north and two potential sites in the south. Determine the number of possible outcomes.

For the first experiment there are 4 possible outcomes A, B, C and D and for the second 2 possible outcomes E and F. Thus the total number of possible outcomes is given by the multiplication rule as 8. Figure 9.2 shows the tree diagram.

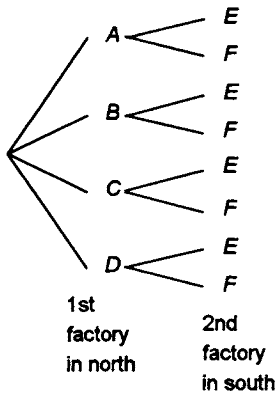


Figure 9.2 *Example*

Permutations

Suppose we had to select two items from a possible three different items A, B, C. The first item can be selected in three ways. Then, since the removal of the first item leaves just two remaining, the second item can be selected in two ways. Thus the selections we can have are:

AB AC BA BC CA CB

Each of the ordered arrangements is known as a *permutation*, each representing the way distinct objects can be arranged.

If there are n ways of selecting the first object, there will be $(n - 1)$ ways of selecting the second object, $(n - 2)$ ways of selecting the third object and $(n - r + 1)$ ways of selecting the rth object. Thus, by the multiplication rule, the total number of different permutations of selecting r objects from n distinct objects is thus:

$$n(n-1)(n-2)\ \dots\ (n-r+1)$$

The number $n(n - 1)(n - 2)\ \dots\ (3)(2)(1)$ is represented by $n!$. The number of permutations of k objects chosen from n distinct objects is represented by ${}^nP_r$ or ${}_nP_r$ or $\binom{n}{r}$ and is thus:

$$ {}^nP_r = n(n-1)(n-2)\ \dots\ (n-r+1) = \frac{n!}{(n-r)!} \qquad [4]$$

r taking values from 0 to n. Note that $0!$ is taken as having the value 1. The number of permutations of n objects chosen from n distinct objects is represented by ${}^nP_n$ or $\binom{n}{n}$ and is thus:

$$ {}^nP_n = n! \qquad [5]$$

Example

In the wiring up of an electronic component there are four assemblies that can be wired up in any order. In how many different ways can the component be wired?

This involves determining the number of permutations of four objects from four. Thus, using equation [5]:

${}^nP_n = n! = 4! = 24$.

> **Example**
>
> How many four-digit numbers can be formed from the digits 0 to 9 if no digit is to be repeated within any one number?
>
> This involves determining the number of permutations of 4 objects from 10. Thus, using equation [4]:
>
> $$^nP_n = \frac{n!}{(n-r)!} = \frac{10!}{(10-4)!}$$
>
> $$= \frac{10\times9\times8\times7\times6\times5\times4\times3\times2\times1}{6\times5\times4\times3\times2\times1} = 5040$$

Combinations

There are often situations where we want to know the number of ways r items can be selected from n objects without being concerned with the order in which the objects are selected. Suppose we had to select two items from a possible three different items A, B, C. The selections, i.e. permutations, we can have are:

AB AC BA BC CA CB

But if we are not concerned with the sequence of the letters then we only have the three ways AB, AC and BC. Such an unordered set is termed a *combination*.

Consider the selection of a combination of r items from n distinct objects. In the selected r items there will be $r!$ permutations (equation [5]) of distinct objects so that the permutation of r items from n contains each group of r items $r!$ times. Since there are $n!/(n-r)!$ different permutations of r items from n we must have:

$$r! \times {}^nC_r = \frac{n!}{(n-r)!}$$

where ${}^nC_r$, ${}_nC_r$ or $\binom{n}{r}$ is used to represent the combination of r items from n. Thus:

$$^nC_r = \frac{n!}{r!(n-r)!} \qquad [6]$$

${}^nC_r$ is often termed a *binomial coefficient*. This is because numbers of this form appear in the expansion of $(x + y)^n$ by the binomial theorem (see Section 7.1.2).

When r items are selected from n distinct objects, $n - r$ items are left. The number of ways of selecting r items from n is given by equation [6] as $n!/r!(n-r)!$. The number of ways of selecting $n - r$ items from n is given by equation [6] as:

$$^nC_{n-r} = \frac{n!}{(n-r)!(n-\{n-r\})!} = \frac{n!}{r!(n-r)!}$$

Thus we can say that there are as many ways of selecting r items from n as selecting $n-r$ objects from n:

$$^nC_r = {^n}C_{n-r} \qquad [7]$$

There is just one combination of n items from n objects. Thus $^nC_n = 1$. If we select 0 items from n, then because equation [6] gives $^nC_0 = n!/0!$ and we take $1/0! = 1$, we have $^nC_0 = 1$. Evidently there are as many ways of selecting none of the items in a set of n as there are of choosing the n objects that are left.

Example

In how many ways can three objects be chosen from a sample of 20?

Using equation [6]: $^{20}C_3 = \frac{20!}{3!17!} = \frac{20 \times 19 \times 18}{1 \times 2 \times 3} = 1140$

Example

If a batch of 20 objects contains 3 with faults and a sample of 5 is chosen, what is the probability of obtaining a sample with (a) 0, (b) 1, (c) 2 faulty items?

The number of ways we choose the sample of 5 items out of 20 is, using equation [6]:

$$^{20}C_5 = \frac{20!}{5!15!}$$

(a) The number of ways we can choose a sample with 0 defective items is the number of ways we choose 5 items from 17 good items and is thus:

$$^{17}C_5 = \frac{17!}{5!12!}$$

Thus the probability of choosing a sample with 0 faulty items is:

$$\text{probability} = \frac{^{17}C_5}{^{20}C_5} = \frac{17!}{5!15!}\frac{5!15!}{20!} = \frac{91}{228}$$

(b) The number of ways we can choose a sample with 1 faulty item and 4 good items, i.e. selecting 1 faulty item from 3 faulty items and 4 good items from 17 good items, is given by the multiplication rule as $^3C_1 \times {^{17}C_4}$. Thus the probability of choosing a sample with 1 faulty item is:

$$\text{probability} = \frac{\dfrac{3!}{1!2!} \times \dfrac{17!}{4! \times 13!}}{\dfrac{20!}{5!15!}} = \frac{35}{76}$$

(c) The number of ways we can choose a sample with 2 faulty items and 3 good items, i.e. selecting 2 faulty items from 3 faulty items and 3 good items from 17 good items, is given by the multiplication rule as ${}^3C_2 \times {}^{17}C_3$. Thus the probability of choosing a sample with 2 faulty items is:

$$\text{probability} = \frac{\dfrac{3!}{2!1!} \times \dfrac{17!}{3! \times 14!}}{\dfrac{20!}{5!15!}} = \frac{5}{38}$$

Conditional probability

The multiplication rule is only valid when the occurrence of one event has no effect upon the probability of the second event occurring. While this can be used in many situations, there are situations where a successful occurrence of the first event affects the probability of occurrence of the second event. Suppose we have 50 objects of which 15 are faulty. What is the probability that the second object selected is faulty given that the first object selected was fault-free? This is a probability problem where the answer depends on the additional knowledge given that the first selection was fault-free. This means that there are less fault-free objects among those remaining for the second selection. Such a problem is said to involve *conditional probability.*

Example

Suppose we have 50 objects of which 15 are faulty. What is the probability that the second object selected is faulty given that the first object selected was fault-free?

Selecting the first object from 50 as fault-free has a probability of 35/50. Because the first object was fault-free we now have 34 fault-free and 15 faulty objects remaining. Now selecting a faulty object from 49 has a probability of 15/49. Using the multiplication rule gives the probability of the first object being fault-free followed by the second object faulty as (35/50)(15/49) = 0.21.

Problems 9.1

1 In a testing period of 1 year, 4 out of 50 of the items tested failed. What is the probability of finding one of the items failing?

2 In a pack of 52 cards there are 4 aces. What is the probability of selecting, at random, an ace from a pack?

3 Testing of a particular item bought for incorporation in a product shows that of 100 items tested, 4 were found to be faulty. What is the probability that one item taken at random will be (a) faulty, (b) free from faults?

4 Resistors manufactured as 10 Ω by a company are tested and 5% are found to have values below 9.5 Ω and 10% above 10.5 Ω. What is the probability that one resistor selected at random will have a resistance between 9.5 Ω and 10.5 Ω?

5 100 integrated circuits are tested and 3 are found to be faulty. What is the probability that one, taken at random, will result in a working circuit?

6 Tests of an electronic product show that 1% have defective integrated circuits alone, 2% have defective connectors alone and 1% have both defective integrated circuits and connectors. What is the probability of one of the products being found to have a (a) defective integrated circuit alone, (b) defective integrated circuit, (c) defective connector, (d) no defects?

7 Cars coming to a junction can turn to the left, to the right or go straight on. If observations indicate that all the possible outcomes are equally likely, determine the probability that a car will (a) go straight on, (b) turn from the straight-on direction.

8 In how many ways can (a) 8 items be selected from 8 distinct objects, (b) 4 items be selected from 7 distinct items, (c) 2 items be selected from 6 distinct items?

9 In how many ways can (a) 2 items be selected from 7 objects, (b) 5 items be selected from 7 objects, (c) 7 items be selected from 7 objects?

10 How many samples of 4 can be taken from 25 items?

11 A batch of 24 components includes 2 that are faulty. If a sample of 2 is taken, what is the probability that it will contain (a) no faulty components, (b) 1 faulty component, (c) 2 faulty components?

12 A batch of 10 components includes 3 that are faulty. If a sample of 2 is taken from the batch, what is the probability that it will contain (a) no faulty components, (b) 2 faulty components?

13 Of 10 items manufactured, 2 are faulty. If a sample of 3 is taken at random, what is the probability it will contain (a) both the faulty items, (b) at least 1 faulty item?

14 A security alarm system is activated and deactivated by keying-in a three-digit number in the proper sequence. What is the total number of possible code combinations if digits may be used more than once?

15 When checking on the computers used in a company it was found that the probability of one having the latest microprocessor was 0.8, the probability of having the latest software 0.6 and the probability of having the latest processor and latest software 0.3. Determine the probability that a computer selected as having the latest software will also have the latest microprocessor.

9.2 Distributions

Key points

Quantities whose variation contains an element of chance are called *random variables*.

Variables which can only assume a number of particular values are called *discrete variables*. Variables which can assume any value in some range are called *continuous variables*.

For example, if we count the number of times per hour that cars pass a particular point then the result will be series of numbers such as 12, 30, 17, etc. The variable is thus discrete. However, if we repeatedly measure the time taken for 100 oscillations of a pendulum then the results will vary as a result of experimental errors and will be a series of values within a range of, say, 20.0 to 21.0 s. The variable can assume any value within that range and so is said to be a continuous variable.

Terms used in industry
In manufacturing, two types of 'variables' are encountered, namely variables and attributes. Generally speaking, variables are quantities which may be quantitatively measured against a calibrated standard, e.g. voltage in volts, mass in kg, temperature in K, etc. Variables which cannot be measured against a set standard, e.g. taste, smell, colour, etc. are termed attributes and are more difficult to control. There are, of course, grey areas; for example, taste is subjective unless broken down into clearly defined chemical constituents and measured chemically against quantitative standards.

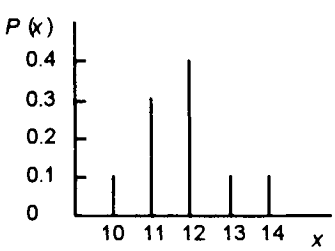


Figure 9.3 *Probability distribution with a discrete variable*

All measurements are affected by random uncertainties and thus repeated measurements will give readings which fluctuate in a random manner from each other. This section considers the statistical approach to such variability of data, dealing with the measures of location and dispersion, i.e. mean, standard deviation and standard error, and the binomial, Poisson and Normal distributions. This statistical approach to variability is especially important in the production environment and the consideration of the variability of measurements made on products.

9.2.1 Probability distributions

Consider the collection of data on the number of cars per hour passing some point and suppose we have the following results:

10, 12, 11, 13, 11, 12, 14, 12, 12, 11

When the discrete variable is sampled 10 times the value 10 appears once. Thus the probability of 10 appearing is 1/10. The value 11 appears three times and so its probability is 3/10, 12 has a probability of 4/10, 13 has the probability 1/10, 14 has the probability 1/10. Figure 9.3 shows how we can represent these probability values as a *probability distribution*.

Consider some experiment in which repeated measurements are made of the time taken for 100 oscillations of a simple pendulum and suppose we have the following results:

20.1, 20.3, 20.8, 20.5, 21.0, 20.8, 20.3, 20.4, 20.7, 20.6,
20.5, 20.7, 20.5, 20.1, 20.6, 20.4, 20.7, 20.5, 20.6, 20.3

With a continuous variable there are an infinite number of values that can occur within a particular range so the probability of one particular value occurring is effectively zero. However, it is meaningful to consider the probability of the variable falling within a particular subinterval. The term *frequency* is used for the number of times a measurement occurs within an interval and the term *relative frequency* or *probability* $P(x)$ for the fraction of the total number of readings in a segment. Thus if we divide the range of the above results into 0.2 intervals, we have:

values >20.0 and ≤20.2 come up twice, thus $P(x) = 2/20$
values >20.2 and ≤20.4 come up five times, thus $P(x) = 5/20$
values >20.4 and ≤20.6 come up seven times, thus $P(x) = 7/20$
values >20.6 and ≤20.8 come up five times, thus $P(x) = 5/20$
values >20.8 and ≤21.0 come up once, thus $P(x) = 1/20$

The probability always has a value less than 1 and the sum of all the probabilities is 1. Figure 9.4 shows how we can represent this graphically. The probability that x lies within a particular interval is thus the height of the rectangle for that strip divided by the sum

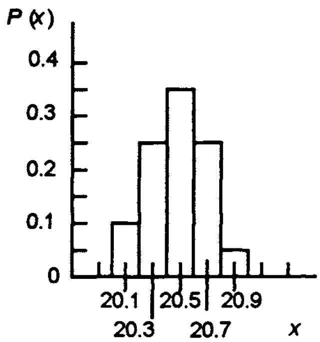


Figure 9.4 *Probability distribution with a continuous variable*

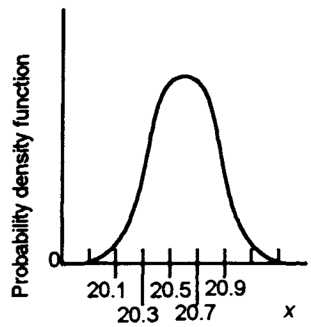


Figure 9.5 *Probability distribution function*

Key point

The *probability density function* is a function that allocates probabilities to all of the range of values that the random variable can take. The probability that the variable will be in any particular interval is obtained by integrating the probability density function over that interval.

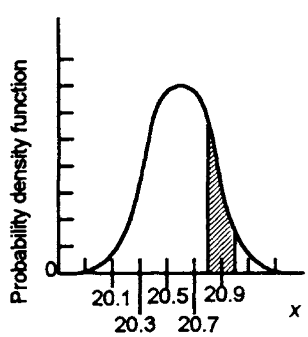


Figure 9.6 *Probability for interval 20.8 to 21.0*

of the heights of all the rectangles. Since each strip has the same width w:

$$\text{probability of } x \text{ being in an interval} = \frac{\text{area of strip}}{\text{total area}} \qquad [8]$$

The histogram shown in Figure 9.4 has a jagged appearance. This is because it represents only a few values. If we had taken a very large number of readings then we could have divided the range into smaller segments and still had an appreciable number of values in each segment. The result of plotting the histogram would now be to give one with a much smoother appearance. When the probability distribution graph is a smooth curve, with the area under the curve scaled to have the value 1, then it is referred to as the *probability density function* $f(x)$ (Figure 9.5). Then equation [8] gives:

probability of x being in the interval $a < x \leq b$

= area under the function of that interval

$$= \int_a^b f(x)\,dx \qquad [9]$$

Consider the probability, with a very large number of readings, of obtaining a value between 20.8 and 21.0 with the probability distribution function shown in Figure 9.6. If we take a segment 20.8 to 21.0 then the area of that segment is the probability. If, say, the area is 0.30, the probability of taking a single measurement and finding it in that interval is 0.30, i.e. the measurement occurs on average 30 times in every 100 values taken.

Example

The following readings, in metres, were made for a measurement of the distance travelled by an object in 10 s. Plot the results as a distribution with segments of width 0.01 m.

13.478, 13.509, 13.502, 13.457, 13.492, 13.512, 13.475, 13.504, 13.473, 13.482, 13.492, 13.500, 13.493, 13.501, 13.472, 13.477

With segments of width 0.01 m we have:

Segment 13.45 to 13.46, frequency 1, so probability 1/16
Segment 13.46 to 13.47, frequency 0, so probability 0
Segment 13.47 to 13.48, frequency 5, so probability 5/16
Segment 13.48 to 13.49, frequency 1, so probability 1/16
Segment 13.49 to 13.50, frequency 4, so probability 4/16
Segment 13.50 to 13.51, frequency 4, so probability 4/16
Segment 13.51 to 13.52, frequency 1, so probability 1/16

Figure 9.7 shows the resulting distribution.

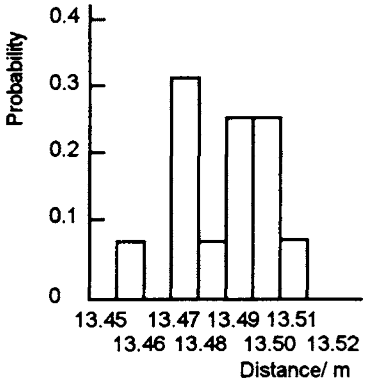


Figure 9.7 *Example*

Key point

The mean is used as a measure of the location of a distribution and the standard deviation its spread.

9.2.2 Measures of location and spread of a distribution

Parameters which can be specified for distributions to give an indication of location and a measure of the dispersion or spread of the distribution about that value are the *mean* for the location and the *standard deviation* for the measure of dispersion.

Mean

The mean value of a set of readings can be obtained in a number of ways, depending on the form with which the data is presented:

- For a list of discrete readings, sum all the readings and divide by the number N of readings, i.e.:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_j}{N} = \frac{\Sigma x_j}{N} \qquad [10]$$

- For a distribution of discrete readings, if we have n_1 readings with value x_1, n_2 readings with value x_2, n_3 readings with value x_3, etc., then the above equation for the mean becomes:

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + n_3 x_3 + \ldots + n_j x_j}{N} \qquad [11]$$

But n_1/N is the relative frequency or probability of value x_1, n_2/N is the relative frequency or probability of value x_2, etc. Thus, to obtain the mean, multiply each reading by its relative frequency or probability P and sum over all the values:

$$\bar{x} = \sum_{j=1}^{n_j} P_j x_j \qquad [12]$$

- For readings presented as a continuous distribution curve, we can consider that we have a discrete-value distribution with very large numbers of very thin segments. Thus if $f(x)$ represents the probability distribution and x the measurement values, the probability that x will lie in a small segment of width δx is $f(x)\delta x$. Thus the rule given above for discrete-value distributions translates into:

$$\bar{x} = \int_{-\infty}^{\infty} x f(x)\,dx \qquad [13]$$

With a very large number of readings, the mean value is taken as being the *true value* about which the random fluctuations occur. The mean value of a probability distribution function is often termed the *expected value*.

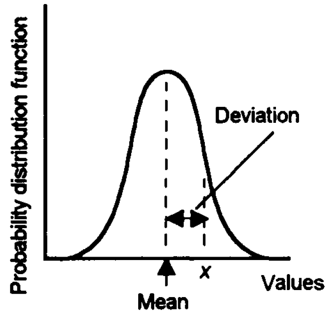


Figure 9.8 *Deviation*

Key point

Accuracy and precision
Consider two marksmen firing 5 shots, from the same gun under the same conditions, at a target with the aim of hitting the central bulls-eye. If the results obtained are as shown in Figure 9.10, then marksman A is accurate in that his shots have a mean which coincides with the bulls-eye. Marksman B has a smaller scatter of results but his mean is not centres on the bulls-eye. His shots have greater precision but are less accurate. Dispersion measurement is thus extremely important when designing manufacturing processes and machining to a set target mean value, i.e. the bulls-eye, with attainable tolerances.

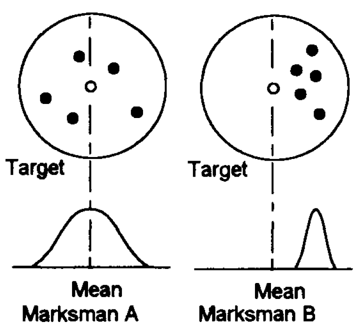


Figure 9.10 *Accuracy and precision*

Key point

The standard deviation σ is the root-mean-square value of the deviations for all the measurements in the distribution.

Standard deviation

Any single reading x in a distribution (Figure 9.8) will deviate from the mean of that distribution by:

$$\text{deviation} = x - \overline{x} \qquad [14]$$

With one distribution we might have a series of values which is widely scattered around the mean while another has readings closely grouped round the mean. Figure 9.9 shows the type of curves that might occur.

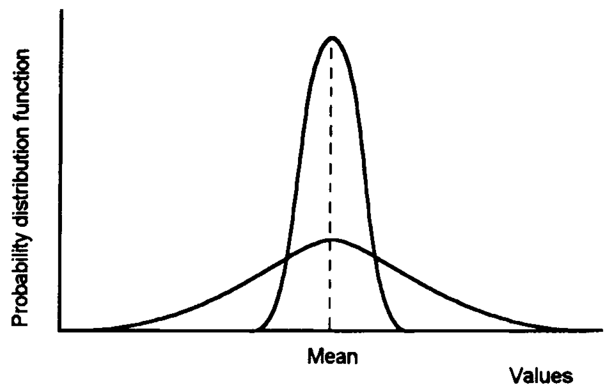


Figure 9.9 *Distributions with different spreads but the same mean*

A measure of the spread of a distribution *cannot* be obtained by taking the mean deviation from the mean, since for every positive value of a deviation there will be a corresponding negative deviation and so the sum will be zero. The measure used is the *standard deviation.*

The standard deviation σ is the root-mean-square value of the deviations for all the measurements in the distribution. The quantity σ^2 is known as the variance of the distribution.

Thus, for a number of discrete values, x_1, x_2, x_3, ..., etc., we can write for the mean value of the sum of the squares of their deviations from the mean of the set of results:

$$\text{sum of squares of deviation} = \frac{\left(x_1 - \overline{x}\right)^2 + \left(x_2 - \overline{x}\right)^2 + \left(x_3 - \overline{x}\right)^2 + \ldots}{N}$$

Hence the mean of the square root of this sum of the squares of the deviations, i.e. the standard deviation, is:

$$\sigma = \sqrt{\frac{\left(\left(x_1 - \overline{x}\right)^2 + \left(x_2 - \overline{x}\right)^2 + \left(x_3 - \overline{x}\right)^2 + \ldots\right)}{N}} \qquad [15]$$

However, we need to distinguish between the standard deviation s of a sample and the standard deviation σ of the entire population of readings that are possible and from which we have only considered a sample (many statistics textbooks adopt the convention of using Greek letters when referring to the entire population and Roman for samples). When we are dealing with a sample we need to write:

$$s = \sqrt{\frac{\left(\left(x_1 - \overline{x}_s\right)^2 + \left(x_2 - \overline{x}_s\right)^2 + \left(x_3 - \overline{x}_s\right)^2 + \ldots\right)}{N-1}} \qquad [16]$$

with $\overline{x}_s$ being the mean of the sample. The reason for using $N - 1$ rather than N is that the root-mean-square of the deviations of the readings in a sample around the sample mean is less than around any other figure. Hence, if the true mean of the entire population were known, the estimate of the standard deviation of the sample data about it would be greater than that about the sample mean. Therefore, by using the sample mean, an underestimate of the population standard deviation is given. This bias can be corrected by using one less than the number of observations in the sample in order to give the sample mean.

For a continuous probability density distribution, since $(n_i/N)\,\delta x$ is the probability for that interval δx, i.e. $f(x)\,\delta x$ where $f(x)$ is the probability distribution function, the standard deviation becomes:

$$\sigma = \int_{-\infty}^{\infty} \left(x - \overline{x}\right)^2 f(x)\,\mathrm{d}x \qquad [17]$$

We can write this equation in a more useful form for calculation:

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x - 2\overline{x}\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x + \overline{x}^2 \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$$

Since the total area under the probability density function curve is 1, the third integral has the value 1 and so the third term is the square of the means. The second integral is the mean. The first integral is the mean value of x^2. Thus:

$$\sigma^2 = \overline{x^2} - \overline{x}^2 \qquad [18]$$

i.e. the mean value of x^2 minus the square of the mean value.

Example

Determine the mean value and the standard deviation of the sample of 10 readings 8, 6, 8, 4, 7, 5, 7, 6, 6, 4.

The mean value is (8 + 6 + 8 + 4 + 7 + 5 + 7 + 6 + 6 + 4)/10 = 6.1.

The standard deviation of the sample can be calculated by considering the deviations of each reading from the mean, these being:

1.9, –0.1, 1.9, –2.1, 1.9, –1.1, 0.9, –0.1, –0.1, –2.1

The squares of these deviations are:

3.61, 0.01, 3.61, 4.41, 3.61, 1.21, 0.81, 0.01, 0.01, 4.41

The sum of these squares is 21.7. If we consider we do not have the entire population but just a sample, then the standard deviation is $\sqrt{(21.7/9)} = 1.6$.

Example

In an experiment involving the counting of the number of events that occurred in equal size time intervals the following data was obtained:

0 events 13 times, 1 event 12 times, 2 events 9 times
3 events 5 times, 4 events once

Determine the mean number of events occurring in the time interval and the standard deviation.

The total number of measurements is 13 + 12 + 9 + 5 + 1 = 40 and so the mean value is $(0 \times 13 + 1 \times 12 + 2 \times 9 + 3 \times 5 + 4 \times 1)/40 = 1.25$.

We have 13 measurements with deviation –1.25, 12 with deviation –0.25, 9 with deviation +0.75, 5 with deviation +1.75 and 1 event with deviation 2.75. We can take the squares of these deviations, sum them and divide by (40 – 1). Hence the standard deviation is 1.1.

Example

A probability density function has $f(x) = 1$ for $0 \leq x < 1$ and elsewhere 0. Determine its mean and standard deviation.

The mean value is given by equation [13] as:

$$\bar{x} = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x\,dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}$$

The standard deviation is given by equation [17]:

$$\sigma^2 = \int_{-\infty}^{\infty}(x-\bar{x})^2 f(x)\,dx = \int_0^1\left(x-\tfrac{1}{2}\right)^2 dx = \int_0^1\left(x^2-x+\tfrac{1}{4}\right) d$$

$$= \left[\frac{x^3}{3}-\frac{x^2}{2}+\frac{x}{4}\right]_0^1 = \tfrac{1}{3}-\tfrac{1}{2}+\tfrac{1}{4} = \tfrac{1}{12} = 0.29$$

Alternatively, using equation [18]:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

since we have:

$$\overline{x^2} = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 x^2\,dx = \left[\frac{x^3}{3}\right]_0^1 = \tfrac{1}{3}$$

then:

$$\sigma^2 = \tfrac{1}{3}-\left(\tfrac{1}{2}\right)^2 = \tfrac{1}{3}-\tfrac{1}{4} = \tfrac{1}{12} = 0.29$$

Standard error of the mean

With a sample set of readings taken from a large population we can determine its mean, but what is generally required is an estimate of the error of that mean from the true value, i.e. the mean of an infinitely large number of readings. We can consider any set of readings as being just a sample taken from the very large set.

Consider one sample of readings with n values being taken: x_1, x_2, x_3, ... x_n. The mean of this sample is:

$$\bar{x} = \frac{1}{n}\sum x_j$$

This mean will have a deviation or error E from the true mean value X of :

$$E = \bar{x} - X.$$

Hence we can write:

$$E = \left(\frac{1}{n}\sum x_j\right) - X = \frac{1}{n}\sum\left(x_j - X\right)$$

If we write e_1 for the error of the first reading from the true mean, e_2 for the error of the second, etc. we obtain:

$$E = \frac{1}{n}\left(e_1 + e_2 + e_3 + \dots e_j\right)$$

Thus:

$$E^2 = \frac{1}{n^2}(e_1^2 + e_2^2 + e_3^2 + \dots + \text{products such as } e_1e_2, \text{ etc.})$$

E is the error from the mean for a single sample of readings. Now, consider a large number of such samples with each set having the same number n of readings. We can write such an equation as above for each sample. If we add together the equations for all the samples and divide by the number of samples considered, we obtain an average value over all the samples of E^2. Thus E is the standard deviation of the means and is known as the *standard error of the means* e_m (more usually the symbol σ). Adding together all the error product terms will give a total value of zero, since as many of the error values will be negative as well as positive. The average of all the Σe_j^2 terms is ne_s^2, where e_s is, what can be termed, the *standard error of the sample*. Thus:

$$e_m = \frac{e_s}{\sqrt{n}}$$

Key point

The standard error of the mean obtained from a sample is the standard deviation of the sample divided by the square root of the sample size.

But how can we obtain a measure of the standard error of the sample? The standard error is measured from the true value X, which is not known. What we can measure is the standard deviation of the sample from its mean value. The best estimate of the standard error for a sample turns out to be the standard deviation s of a sample when we define it as:

$$s^2 = \frac{1}{n-1} \Sigma(x_j - \overline{x})^2$$

i.e. with a denominator of $N - 1$, rather than just N. Thus the best estimate of the standard error of the mean can be written as:

$$\text{standard error of the mean} = \frac{s}{\sqrt{n}} \qquad [19]$$

Example

Measurements are to be made of the percentage of an element in a chemical by making measurements on a number of samples. The standard deviation of any one sample is found to be 2%. How many measurements must be made to give a standard error of 0.5% in the estimated percentage of the element.

If n measurements are made, then the standard error of the sample mean is given by equation [19] and so:

$$n = \frac{2^2}{0.5^2} = 16$$

9.2.3 Common distributions

There are three basic forms of distribution which are found to represent many forms of distributions commonly encountered in engineering and science. These are the binomial distribution, the

Poisson distribution and the normal distribution (sometimes called the Gaussian distribution). Binomial distributions are often approximated by the Poisson distribution. The normal distribution is a model widely used for experimental measurements when there are random errors.

Binomial distribution

In the tossing of a single coin the result is either heads or tails uppermost. We can consider this as an example of an experiment where the results might be termed as either success or failure, one result being the complement of the other. If the probability of succeeding is p then the probability of failing is $1 - p$. Such a form of experiment is termed a *Bernoulli trial.*

Suppose the trial is the throwing of a die with a 6 uppermost being success. The probability of obtaining a 6 as the result of one toss of the die is 1/6 and the probability of not obtaining a 6 is 5/6. Suppose we toss the die n times. The probability of obtaining no 6s in any of the trials is given by the product rule as $(5/6)^n$. The probability of obtaining one 6 in, say, just the first trial out of the n is $(5/6)^{n-1}$ (1/6). But we could have obtained the one 6 in any one of the n trials. Thus the probability of one 6 is $n(5/6)^{n-1}$ (1/6). The probability of obtaining two 6s in, say, just the first two trials is $(5/6)^{n-2}(1/6)^2$. But these two 6s may occur in the n trials in a number of combinations $n!/2!(n - 2)$ (see Section 9.1.2). Thus the probability of two 6s in n trials is $[n!/2!(n - 2)](5/6)^{n-2}(1/6)^2$. We can continue this for three 6s, 4s, etc.

In general, if we have n independent Bernoulli trials, each with a success probability p, and of those n trials k give successes, and $(n - k)$ failures, the probability of this occurring is given by the product rule as:

Key point

The characteristics of a variable that gives a binomial distribution are that the experiment consists of n identical trials, there are two possible complementary outcomes, success or failure, for each trial and the probability of a success is the same for each trial, the trials are independent and the distribution variable is the number of successes in n trials.

$$^nC_k p^k(1-p)^{n-k} = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \qquad [20]$$

This is termed the *binomial distribution*. This term is used because, for $k = 0, 1, 2, 3, \ldots n$, the values of the probabilities are the successive terms of the binomial expansion of $[(1 - p) + p]^n$.

For a single Bernoulli trial of a random variable x with probability of success p, the mean value is p. The standard deviation is given by equation [18] as:

$$\sigma^2 = \overline{x^2} - \overline{x}^2 = p^2 - p = p(1-p)$$

For n such trials:

$$\text{mean value} = np \qquad [21]$$

$$\text{standard deviation} = \sqrt{np(1-p)} \qquad [22]$$

Key point

In the example, if the probability of a sale is p and the probability of no sale is q, then $q + p = 1$. We are concerned with 6 enquiries and so if we consider the binomial expansion of $(q + p)^6$ we have:

$$(q + p)^6 = q^6 + 6q^5p + 15q^4p^2 + 20q^3p^3 + 15q^2p^4 + 6qp^5 + p^6$$

Each successive term in the expansion gives the probability of 0, 1, 2, 3, 4, 5 or 6 sales. Thus, with $p = 0.3$ we have $q = 0.7$ and so the probability of 0 sales is $0.7^6 = 0.118$.

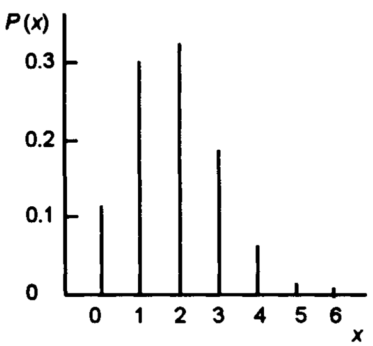


Figure 9.11 *Example*

Example

The probability that an enquiry from a potential customer will lead to a sale is 0.30. Determine the probabilities that among six enquiries there will be 0, 1, 2, 3, 4, 5, 6 sales.

Using equation [20]:

The probability of 0 is:

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{6!}{0!6!}(0.30)^0(0.70)^6 = 0.118$$

The probability of 1 is:

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{6!}{1!5!}(0.30)^1(0.70)^2 = 0.303$$

The probability of 2 is:

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{6!}{2!4!}(0.30)^2(0.70)^4 = 0.324$$

The probability of 3 is:

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{6!}{3!3!}(0.30)^3(0.70)^3 = 0.185$$

The probability of 4 is:

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{6!}{4!2!}(0.30)^4(0.70)^2 = 0.060$$

The probability of 5 is:

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{6!}{5!1!}(0.30)^5(0.70)^1 = 0.010$$

The probability of 6 is:

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{6!}{6!0!}(0.30)^6(0.70)^0 = 0.001$$

Figure 9.11 shows the distribution.

Poisson distribution

The Poisson distribution for a variable λ is:

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad [23]$$

for $k = 0, 1, 2, 3$, etc. The mean of this distribution is λ and the standard deviation is $\sqrt{\lambda}$. When the number n of trials is very large and the probability p small, e.g. $n > 25$ and $p < 0.1$, binomial

Key point

Approximating the binomial distribution to the Poisson distribution

If p is the possibility of an event occurring and q the possibility that it does not occur, then $q + p = 1$ and so if we consider n samples, we have $(q + p)^n = 1$ and the binomial expression gives:

$$(q+p)^n = q^n + nq^{n-1}p + \frac{n(n-1)q^{n-2}p^2}{2!} + \frac{n(n-1)(n-2)q^{n-3}p^3}{3!} + \ldots$$

If p is small then $q = 1 - p \approx 1$ and with n large the first few terms have $n - 1$ approximating to n, $n - 2$ to n, etc. The binomial expression can thus be approximated to:

$$1^n = 1^n + n1^{n-1}p + \frac{nn1^{n-2}p^2}{2!} + \frac{nnn1^{n-3}p^3}{3!} + \ldots$$

$$1^n = 1 + np + \frac{n^2p^2}{2!} + \frac{n^3p^3}{3!} + \ldots$$

If we let $np = \lambda$ then:

$$1^n = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots$$

But this is the series for e^{λ} (see Table 7.1) and so $1^n = e^{\lambda}$. We can thus write the binomial expression as:

$$(q+p)^n = e^{-\lambda}\left[1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\right]$$

and so the terms in the expression are given by equation [23].

probabilities are often approximated by the *Poisson distribution*. Thus, since the mean of the binomial distribution is np (equation [21]) and the standard deviation (equation [22]) approximates to $\sqrt{np}$ when p is small, we can consider λ to represent np. Thus λ can be considered to represent the average number of successes per unit time or unit length or some other parameter.

Example

2% of the output per month of a mass produced product have faults. What is the probability that of a sample of 400 taken that 5 will have faults?

Assuming the Poisson distribution, we have $\lambda = np = 400 \times 0.02 = 8$ and so equation [23] gives for $k = 5$:

$$P(5) = \frac{8^5\,e^{-8}}{5!} = 0.093$$

Example

The output from a CNC machine is inspected by taking samples of 60 items. If the probability of a defective item is 0.0015, determine the probability of the sample having (a) two defective items, (b) more than two defective items.

(a) We have $n = 60$ and $p = 0.0015$. Thus, assuming a Poisson distribution, we have $\lambda = np = 60 \times 0.0015 = 0.09$ and so equation [23] gives the probability of two defective items as $(0.092 \times e^{-0.09})/2! = 3.7 \times 10^{-3}$ or 0.37%.

(b) The probability of there being more than two defective items is $1 - \{P(0) + P(1) + P(2)\} = 1 - e^{-\lambda}\{1 + \lambda + \lambda^2/2!\} = 1 - e^{-0.09}\{1 + 0.09 + 4.5 \times 10^{-3}\} = 1.36 \times 10^{-4}$ or 0.01%.

Example

There is a 1.5% probability that a machine will produce a faulty component. What is the probability that there will be at least 2 faulty items in a batch of 100?

Assuming the Poisson distribution can be used, we have $\lambda = np = 100 \times 0.015 = 1.5$ and so the probability of at least 2 faulty items will be:

$$P(\geq 2) = 1 - P(0) - P(1) = 1 - \frac{1.5^0}{0!}\,e^{-1.5} - \frac{1.5^1}{1!}\,e^{-1.5} = 0.442$$

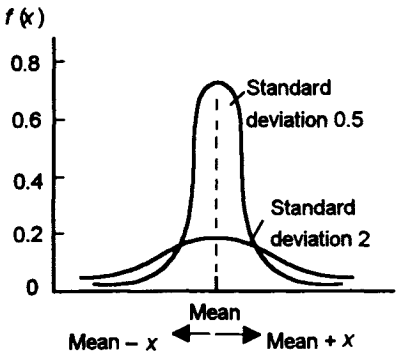


Figure 9.12 *Normal distribution*

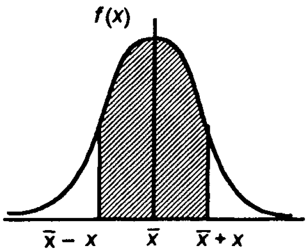


Figure 9.13 *Values within + or − x of the mean*

Key point

The normal distribution is standardised so that we can compare measurements from samples where their units of measure differ, their means differ and their standard deviations may differ, e.g. when comparing similar machining processes for geometrically similar components running off different tolerances.

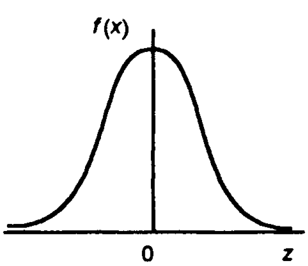


Figure 9.14 *Standard normal distribution*

Normal distribution

A particular form of distribution, known as the *normal distribution* or *Gaussian distribution*, is very widely used and works well as a model for experimental measurements when there are random errors. This form of distribution has a characteristic bell shape (Figure 9.12). It is symmetric about its mean value, having its maximum value at that point, and tends rapidly to zero as x increases or decreases from the mean. It can be completely described in terms of its mean and its standard deviation. The following equation describes how the values are distributed about the mean:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\bar{x})^2/2\sigma^2} \qquad [24]$$

The fraction of the total number of values that lies between $-x$ and $+x$ from the mean is the fraction of the total area under the curve that lies between those ordinates (Figure 9.13). We can obtain areas under the curve by integration.

To save the labour of carrying out the integration, the results have been calculated and are available in tables. As the form of the graph depends on the value of the standard deviation, as illustrated in Figure 9.12, the area depends on the value of the standard deviation σ. In order not to give tables of the areas for different values of x for each value of σ, the distribution is considered in terms of the value of:

$$(x - \overline{x})/\sigma$$

this commonly being designated by the symbol z, and areas tabulated against this quantity. z is known as the *standard normal random variable* and the distributions obtained with this as the variable are termed the *standard normal distribution* (Figure 9.14). Any other normal random variable can be obtained from the standard normal random variable by multiplying by the required standard deviation and adding the mean, i.e.

$$x = \sigma z + \overline{x}$$

Table 9.1 shows examples of the type of data given in tables for z.

When we have:

$$x - \overline{x} = 1\sigma$$

then $z = 1.0$ and the area between the ordinate at the mean and the ordinate at 1σ as a fraction of the total area is 0.341 3. The area within $\pm1\sigma$ of the mean is thus the fraction 0.681 6 of the total area under the curve, i.e. 68.16%. This means that the chance of a value being within $\pm1\sigma$ of the mean is 68.16%, i.e. roughly two-thirds of the values.

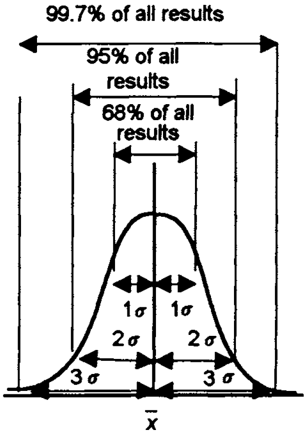


Figure 9.15 *Normal distribution*

Table 9.1 *Areas under normal curve*

z	Area from mean	z	Area from mean
0	0.000 0	1.6	0.445 2
0.2	0.079 3	1.8	0.464 1
0.4	0.155 5	2.0	0.477 2
0.6	0.225 7	2.2	0.486 1
0.8	0.288 1	2.4	0.491 8
1.0	0.341 3	2.6	0.495 3
1.2	0.384 9	2.8	0.497 4
1.4	0.419 2	3.0	0.498 7

Key point

Table 9.1 is only a rough version of the more detailed tables that are needed in statistical process control. Detailed tables are readily available.

When we have:

$$x - \overline{x} = 2\sigma$$

then $z = 2.0$ and the area between the ordinate at the mean and the ordinate at 1σ as a fraction of the total area is 0.477 2. The area within $\pm 2\sigma$ of the mean is thus the fraction 0.954 4 of the total area under the curve, i.e. 95.44%. This means that the chance of a value being within $\pm 2\sigma$ of the mean is 95.44%.

When we have:

$$x - \overline{x} = 3\sigma$$

then $z = 3.0$ and the area between the ordinate at the mean and the ordinate at 3σ as a fraction of the total area is 0.498 7. The area within $\pm 1\sigma$ of the mean is thus the fraction 0.997 4 of the total area, i.e. 99.74%. This means that the chance of a reading being within $\pm 3\sigma$ of the mean is 99.74%. Thus, virtually all the readings will lie within $\pm 3\sigma$ of the mean. Figure 9.15 illustrates the above.

Example

Measurements are made of the tensile strengths of samples taken from a batch of steel sheet. The mean value of the strength is 800 MPa and it is observed that 8% of the samples give values that are below an acceptable level of 760 MPa. What is the standard deviation of the distribution if it is assumed to be normal?

This means that an area from the mean of 0.50 − 0.08 = 0.42. To the accuracy given in Table 9.1, this occurs when $z = 1.4$ (Figure 9.16). Thus,

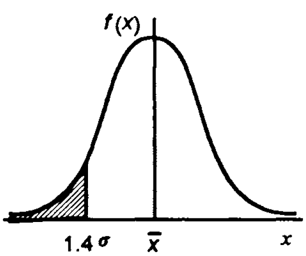


Figure 9.16 *Example*

$$(x - \overline{x})/\sigma = (760 - 800)/\sigma = 1.$$

and so the standard deviation is 29 MPa. In the above analysis, it was assumed that the mean given was the true value or a good enough approximation.

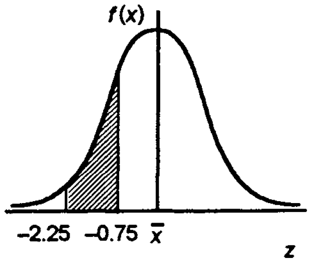


Figure 9.17 *Example*

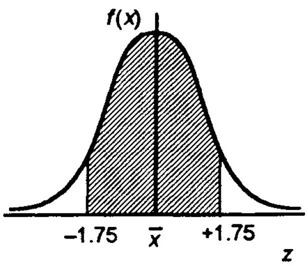


Figure 9.18 *Example*

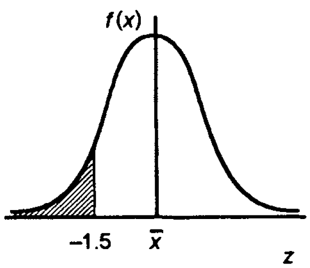


Figure 9.19 *Example*

Example

A pharmaceutical manufacturer produces tablets having a mean mass of 4.0 g and a standard deviation of 0.2 g. Assuming that the masses are normally distributed and that a table is chosen at random, determine the probability that it will (a) have a mass between 3.55 and 3.85 g, (b) will differ from the mean by less than 0.35 g, and (c) determine the number that might be expected to have a mass less than 3.7 g in a carton of 400.

(a) The probability of tablets having masses between 3.55 g and 3.85 g is the area between the normal distribution with ordinates of these masses (Figure 9.17). We have $z_1 = (3.55 - 4.0)/0.2 = -2.25$ and $z_2 = (3.85 - 4.0)/0.2 = -0.75$. Table 9.1, or better tables, gives, approximately, (area between mean and 2.25) − (area between mean and 0.75) = 0.4878 − 0.2734 = 0.2144 or about 21.4%.
(b) To determine the probability of tables differing from the mean of 4.0 g by less than 0.35 g we consider the area between the ordinates for masses between −3.65 g and 4.35 g. These give z values of −1.75 and +1.75 and the area is thus as indicated in Figure 9.18. This is 2 × 0.4599 = 0.9198 or about 92%.
(c) The probability of a mass less than 3.7 g is for a z value less than (3.7 − 4.0)/0.2 = −1.5 (Figure 9.19). The area within +1.5 and −1.5 of the mean is 2 × 0.433 = 0.866. The total area under the curve is 1 and so the area outside these limits is 1 − 0.866 = 0.134. Just half of this area will be for less than 3.7 g and so the area is 0.134. For 400 tablets this means 0.134 × 400 or about 27 tablets.

Problems 9.2

1 Determine the mean value of a variable which can have the discrete values of 2, 3, 4 and 5 and for which 2 occurs twice, 3 occurs three times, 4 occurs three times and 5 occurs once.
2 The probability density function of a random variable x is given by ½x for $0 \le x < 2$ and 0 for all other values. Determine the mean value of the variable.
3 Determine the mean and the standard deviation for the following data: 10, 20, 30, 40, 50.

4 Determine the standard deviation of the resistance values for a sample of 12 resistors taken from a batch if the values are: 98, 95, 109, 99, 102, 99, 106, 96, 101, 108, 94, 102 Ω.

5 Determine the standard deviation of the six values: 1.3, 1.4, 0.8, 0.9, 1.2, 1,0.

6 Determine the standard deviation of the probability distribution function $f(x) = 2x$ for $0 \leq x < 1$ and 0 elsewhere.

7 The following are the results of 100 measurements of the times for 50 oscillations of a simple pendulum:

Between 58.5 and 61.5 s, 2 measurements
Between 61.5 and 64.5 s, 6 measurements
Between 64.5 and 67.5 s, 22 measurements
Between 67.5 and 70.5 s, 32 measurements
Between 70.5 and 73.5 s, 28 measurements
Between 73.5 and 76.5 s, 8 measurements
Between 76.5 and 79.5 s, 2 measurements

(a) Determine the relative frequencies of each segment.
(b) Determine the mean and the standard deviation.

8 A random sample of 25 items is taken and found to have a standard deviation of 2.0. (a) What is the standard error of the sample? (b) What sample size would have been required if a standard error of 0.5 was acceptable?

9 It has been found that 10% of the screws produced are defective. Determine the probabilities that a random sample of 20 will contain 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 defectives.

10 The probability that any one item from a production line will be accepted is 0.70. What is the probability that when 5 items are randomly selected that there will be 2 unacceptable items?

11 Packets are filled automatically on a production line and, from past experience, 2% of them are expected to be underweight. If an inspector takes a random sample of 10, what will be the probability that (a) 0, (b) 1 of the packets will be underweight?

12 1% of the resistors produced by a factory are faulty. If a sample of 100 is randomly taken, what is the probability of the sample containing no faulty resistors?

13 The probability of a mass-produced item being faulty has been determined to be 0.10. What are the probabilities that a random sample of 50 will contain 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 faulty items?

14 A product is guaranteed not to contain more than 2% that are outside the specified tolerances. In a random sample of 10, what is the probability of getting 2 or more outside the specified tolerances?

15 A large consignment of resistors is known to have 1% outside the specified tolerances. What would be the expected number of resistors outside the specified tolerances in a batch of 10 000 and the standard deviation?

16 The number of cars that enter a car park follows a Poisson distribution with a mean of 4. If the car park can accommodate 12 cars, determine the probability that the car park is filled up by the end of the first hour it is open.

17 On average six of the cars coming per day off a production line have faults. What is the probability that four faulty cars will come off the line in one day?

18 The number of breakdowns per month for a machine averages 1.8. Determine the probability that the machine will function for a month with only one breakdown.

19 Measurements of the resistances of resistors in a batch gave a mean of 12 Ω with a standard deviation of 2 Ω. If the resistances can be assumed to have a normal distribution about this mean, how many from a batch of 300 resistors are likely to have resistances more than 15 Ω?

20 Measurements made of the lengths of components as they come off the production line have a mean value of 12 mm with a standard deviation of 2 mm. If a normal distribution can be assumed, in a sample of 100 how many might be expected to have (a) lengths of 15 mm or more, (b) lengths between 13.7 and 16.1 mm?

21 Measurements of the times taken for workers to complete a particular job have a mean of 29 minutes and a standard deviation of 2.5. Assuming a normal distribution, what percentage of the times will be (a) between 31 and 32 minutes, (b) less than 26 minutes?

22 Inspection of the lengths of components yields a normal distribution with a mean of 102 mm and a standard deviation of 1.5 mm. Determine the probability that if a component is selected at random it will have a length (a) less than 100 mm, (b) more than 104 mm, (c) between 100 and 104 mm.

23 A machine makes resistors with a mean value of 50 Ω and a standard deviation of 2 Ω. Assuming a normal distribution, what limits should be used on the values of the resistance if there are to be not more than 1 reject in 1000.

24 A series of measurements was made of the periodic time of a simple pendulum and gave a mean of 1.23 s with a standard deviation of 0.01 s. What is the chance that, when a measurement is made, it will lie between 1.23 and 1.24 s?

25 The measured resistance per metre of samples of a wire have a mean resistance of 0.13 Ω and a standard deviation of 0.005 W. Determine the probability that a randomly selected wire will have a resistance per metre of between 0.12 and 0.14 Ω.

26 A set of measurements has a mean of 10 and a standard deviation of 5. Determine the probability that a measurement will lie between 12 and 15.

27 A set of measurements has a normal distribution with a mean of 10 and a standard deviation of 2.1. Determine the probability of a reading having a value (a) greater than 11 and (b) between 7.6 and 12.2.

9.3 Experimental errors

Experimental *error* is defined as the difference between the result of a measurement and the true value:

$$\text{error} = \text{measured value} - \text{true value} \qquad [25]$$

Errors can arise from such causes as instrument imperfections, human imprecision in making measurements and random fluctuations in the quantity being measured. This section is a brief consideration of the estimation of errors and their determination in a quantity which is a function of more than one measured variable.

With measurements made with an instrument, errors can arise from fluctuations in readings of the instrument scale due to perhaps the settings not being exactly reproducible and operating errors because of human imprecision in making the observations. The term *random error* is used for those errors which vary in a random manner between successive readings of the same quantity. The term *systematic error* is used for errors which do not vary from one reading to another, e.g. those arising from a wrongly set zero. Random errors can be determined and minimised by the use of statistical analysis, systematic errors require the use of a different instrument or measurement technique to establish them.

With random errors, repeated measurements give a distribution of values. This can be generally assumed to be a normal distribution. The standard error of the mean of the experimental values can be estimated from the spread of the values and it is this which is generally quoted as the error, there being a 68% probability that the mean will lie within plus or minus one standard error of the true mean. Note that the standard error does not represent the maximum possible error. Indeed there is a 32% probability of the mean being outside the plus and minus standard error interval.

With just a single measurement, say the measurement of a temperature by means of a thermometer, the error is generally quoted as being plus or minus one-half of the smallest scale division. This, termed the *reading error*, is then taken as an estimate of the standard deviation that would occur for that measurement if it had been repeated many times.

Key points

Random errors are those errors which vary in a random manner between successive readings of the same quantity. Random errors can be determined and minimised by the use of statistical analysis.

Systematic errors are those errors which do not vary from one reading to another, e.g. those arising from a wrongly set zero. Systematic errors require the use of a different instrument or measurement technique to establish them.

Example

A rule used for the measurement of a length has scale readings every 1 mm. Estimate the error to be quoted when the rule is used to make a single measurement of a length.

The error is quoted as ±0.5 mm.

Example

Measurements of the tensile strengths of test pieces taken from a batch of incoming material gave the following results: 40, 42, 39, 41, 45, 40, 41, 43, 45, 46 MPa. Determine the mean tensile strength and its error.

The mean is given by equation [10] as (40 + 42 + 39 + 41 + 45 + 40 + 41 + 43 + 45 + 44)/10 = 42. The standard deviation can be calculated by the use of equation [12]. The deviations from the mean are –2, 0, –3, –1, 3, –1, 1, 3, 4 and their squares are 4, 0, 9, 1, 9, 1, 1, 9, 16. The standard deviation is thus √[50/(10 – 1)] = 2.4 MPa. The standard error is thus 2.4/√10 = 0.8 MPa. Thus the result can be quoted as 42 ± 0.8 MPa.

Statistical errors

In addition to measurement errors arising from the use of instruments, there are what might be termed *statistical errors*. These are not due to any errors arising from an instrument but from statistical fluctuations in the quantity being measured, e.g. the count rate of radioactive materials. The observed values here are distributed about their mean in a Poisson distribution and so the standard deviation is the square root of the mean value.

Example

In an experiment, the number of alpha particles emitted over a fixed period of time is measured as 4206. Determine the standard deviation of the count.

Assuming the count follows a Poisson distribution, the standard deviation will be the square root of 4206 and so 65. Thus the count can be recorded as 4206 ± 65.

9.3.1 Combining errors

An experiment might require several quantities to be measured and then the values inserted into an equation so that the required variable can be calculated. For example, a determination of the density of a material might involve a determination of its mass and volume, the density then being calculated from mass/volume. If the mass and volume each have errors, how do we combine these errors in order to determine the error in the density?

Consider a variable Z which is to be determined from sets of measurements of A and B and for which we have the relationship $Z = A + B$. If we have A with an error ΔA and B with an error ΔB

then we might consider that $Z + \Delta Z = A + \Delta A + B + \Delta B$ and so we should have

$$\Delta Z = \Delta A + \Delta B \tag{26}$$

i.e. the error in Z is the sum of the errors in A and B. However, this ignores the fact that the error is the standard error and so is just the value at which there is a probability of ±68% that the mean value for A or B will be within that amount of the true mean. If we consider the set of measurements that were used to obtain the mean value of A and its standard error and the set of measurements to obtain the mean value of B and its standard error and consider the adding together of individual measurements of A and B then we can write $\Delta Z = \Delta A + \Delta B$ for each pair of measurements. Squaring this gives:

$$\Delta Z^2 = (\Delta A + \Delta B)^2 = (\Delta A)^2 + (\Delta B)^2 + 2\,\Delta A\,\Delta B$$

We can write such an equation for each of the possible combinations of measurements of A and B. If we add together all the possible equations and divide by the number of such equations, we would expect the $2\,\Delta A\,\Delta B$ terms to cancel out since there will be as many situations with it having a negative value as a positive value. Thus the equation we should use to find the error in Z is:

$$\Delta Z^2 = (\Delta A)^2 + (\Delta B)^2 \tag{27}$$

The same equation is obtained for $Z = A - B$.

Now consider the error in Z when $Z = AB$. As before, we might argue that $Z + \Delta Z = (A + \Delta A)(B + \Delta B)$ and so $\Delta Z = B\,\Delta A + A\,\Delta B$, if we ignore as insignificant the $\Delta A\,\Delta B$ term. Hence:

$$\frac{\Delta Z}{Z} = \frac{B\,\Delta A + A\,\Delta B}{AB} = \frac{\Delta A}{A} + \frac{\Delta B}{B} \tag{28}$$

i.e. the fractional error in Z is the sum of the fractional errors in A and B or the percentage error in Z is the sum of the percentage errors in A and B. However, this ignores the fact that the error is the standard error and so is just the value at which there is a probability of ±68% that the mean value for A or B will be within that amount of the true mean. If we consider the set of measurements that were used to obtain the mean value of A and its standard error and the set of measurements to obtain the mean value of B and its standard error and use equation [28] for each such combination, then we can write:

$$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta A}{A} + \frac{\Delta B}{B}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2 + 2\left(\frac{\Delta A}{A}\right)\left(\frac{\Delta B}{B}\right)$$

We can write such an equation for each of the possible combinations of measurements of A and B. If we add together all the possible equations and divide by the number of such equations, the $2(\Delta A/A)(\Delta B/B)$ terms cancel out since there will be as many

situations with it having a negative value as a positive value. Thus the equation we should use to find the error in Z is:

$$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2 \qquad [29]$$

The same equation is obtained for $Z = A/B$. If $Z = A^2$ then this is just the product of A and A and so equation [29] gives $(\Delta Z/Z)^2 = 2(\Delta A/A)^2$. Thus for $Z = A^n$ we have:

$$\left(\frac{\Delta Z}{Z}\right)^2 = n\left(\frac{\Delta A}{A}\right)^2 \qquad [30]$$

In all the above discussion it was assumed that the mean value of Z was given when the mean values of A and B were used in the defining equation.

Example

The resistance of a resistor is determined from measurements of the potential difference across it and the current through it. If the potential difference has been measured as 2.1 ± 0.2 V and the current as 0.25 ± 0.01 A, what is the resistance and its error?

The mean resistance is 2.1/0.25 = 8.4 Ω. The fractional error in the potential difference is 0.2/2.1 = 0.095 and the fractional error in the current is 0.01/0.25 = 0.04. Hence the fractional error in the resistance is $\sqrt{(0.095^2 + 0.04^2)}$ = 0.10. Thus the resistance is 8.4 ± 0.9 Ω.

Example

If $g = 4\pi^2 L/T^2$ and L has been measured as 1.000 ± 0.005 m and T as 2.0 ± 0.1 s, determine g and its error.

The mean value of g is $4\pi^2(1.000)/2.0^2$ = 9.87 m/s². The fractional error in L is 0.005 m and that in T is 0.05 s. Thus:

(fractional error in g)$^2 = 0.005^2 + 2 \times 0.05^2$

Thus the fractional error in g is 0.071 and so g = 9.87 ± 0.7 m/s².

Problems 9.3

1 Determine the mean value and standard error for the measured diameter of a wire if it is measured at a number of points and gave the following results: 2.11, 2.05, 2.15, 2.12, 2.16, 2.14, 2.16, 2.17, 2.13, 2.15 mm.

2 An ammeter has a scale with graduations at intervals of 0.1 A. Give an estimate of the standard deviation.

3 Determine the mean and the standard error for the resistance of a resistor if repeated measurements gave 51.1, 51.2, 51.0, 51.4, 50.9 Ω.

4 In an experiment the number of gamma rays emitted over a fixed period of time is measured as 5210. Determine the standard deviation of the count.

5 How big a count should be made of the gamma radiation emitted from a radioactive material if the percentage error should be less than 1%?

6 Repeated measurements of the voltage necessary to cause the breakdown of a dielectric gave the results 38.9, 39.3, 38.6, 38.8, 38.8, 39.0, 38.7, 39.4, 39.7, 38.4, 39.0, 39.1, 39.1, 39.2 kV. Determine the mean value and the standard error of the mean.

7 Determine the mean value and error for Z when (a) $Z = A - B$, (b) $Z = 2AB$, (c) $Z = A^3$, (d) $Z = B/A$ if $A = 100 \pm 3$ and $B = 50 \pm 2$.

8 The resistivity of a wire is determined from measurements of the resistance R, diameter d and length L. If the resistivity is RA/L, where A is the cross-sectional area, which measurement requires determining to the greatest accuracy if it is not to contribute the most to the overall error in the resistivity?

9 The cross-sectional area of a wire is determined from a measurement of the diameter. If the diameter measurement gives 2.5 ± 0.1 mm, determine the area of the wire and its error.

10 Determine the mean value and error for Z when (a) $Z = A + B$, (b) $Z = AB$, (c) $Z = A/B$ if $A = 100 \pm 3$ and $B = 50 \pm 2$.

Solutions to problems

Chapter 1

1.1

1 (a) 3, (b) 5
2 (a) 0, (b) 6
3 (a)
4 (a) 2, (b) 1
5 (a) 0, (b) 2
6 $v = 0$ for $0 \le t \le 2$, $v = 10$ V for $2 \le t$
7 See Figure S.1
8 (a) 2.01 s, (b) 6.34 s
9 (a) 2 m/s, (b) 7 m/s
10 (a) $3x + 1$, (b) $2x + 2$, (c) $2x + 1$
11 (a) $x^2 + 3x + 1$, (b) $9x^2 + 3$, (c) $3x^2 + 3$, (d) $x^2 - 3x - 1$, (e) $9x^2 + 12x + 5$

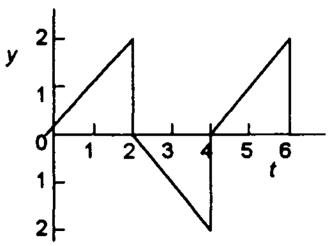


Figure S.1

1.2

1 (a) Straight line through origin, (b) straight line not through origin, (c) not straight line, (d) no straight line
2 (a) $i = 0.5t + 2$ amps, (b) $e = 1.2 \times 10^{-3}\, L$ m

1.3

1 (a) −3.2, 1.2, (b) −2.6, −0.38, (c) 3.8, −0.41, (d) no real roots
2 38.2°C, 261.8°C
3 0.93 m, −0.11m
4 $y = \dfrac{mgL^3}{48EI} \pm \sqrt{\left(\dfrac{mgL^3}{48EI}\right)^2 + 2h\left(\dfrac{mgL^3}{48EI}\right)}$
5 18.8 m or 1.6 m
6 5.41 cm and 8.41 cm

1.4

1 (a) $\frac{1}{5}(x+3)$, (b) $x-4$, (c) $\sqrt[3]{x}$, (d) $\sqrt[3]{\frac{1}{2}(x+1)}$
2 No, but if domain restricted to $x \ge 0$ then yes
3 (a) $x \ge 1,\ 1 + \sqrt{x}$, (b) $x \ge -1,\ -1 + \sqrt{x+4}$

1.5

1 Amplitude = 5, phase angle = +30° leading
2 Amplitude = 4, angular frequency = 3 rad/s
3 (a) 2, 1.27 s, 1 rad, (b) 6, 2.09 s, 0, (c) 5, 9.4 s, 0.33 rad, (d) 2, 6.28 s, –0.6 rad lagging
4 (a) 6, π, 1, (b) 2, $2\pi/9$, 0, (c) 6, 5π, –0.2, (d) 2, 2π, –0.2, (e) 6, $\pi/2$, $\pi/8$, (f) 0.5, 2π, $-\pi/6$
5 40, 20 Hz
6 (a) 0.87 V, (b) –0.87 V
7 (a) 100 mA, (b) 100 Hz, (c) 0.25 rad or 14.3° lagging
8 (a) 12 V, (b) 50 Hz, (c) 0.5 rad or 28.6° lagging
9 16.2 V
10 As given in the problem
11 101.3°, 355.4°
12 (a) $\sqrt{41} \sin(\theta - 5.61)$, (b) $\sqrt{41} \cos(\theta + 5.39)$
13 $R = W(1 + \mu^2)^{1/2}$, $\tan \beta = 1/\mu$
14 (a) $5 \sin(\omega t + 0.927)$, (b) $8.63 \sin(\omega t - 1.01)$, (c) $4.91 \sin(\omega t + 4.13)$ or $4.91 (\omega t - 2.15)$ for $-\pi \le a \le \pi$
15 $5.83 \sin(\theta + 59.03)$
16 (a) $15 \sin(\omega t - 0.64)$, (b) $8.06 \sin(\omega t + 2.62)$, (c) $6.71 \sin(\omega t - 2.03)$
17 $22.4 \sin(\omega t + 1.11)$ mA
18 $6.81 \sin(\omega t + 0.147)$ V
19 $126.2\sqrt{2} \sin(\omega t - 0.071)$ V
20 $8.72\sqrt{2} \sin(\omega t - 9.639)$ A
21 (a) 0.83, (b) 1.47, (c) 0.67, (d) –0.41

1.6

1 (a) The negative sign, (b) N_0, (c) λ small
2 (a) The power is positive, (b) L_0, (c) α high means a high expansion
3 (a) 0, (b) 3 A
4 (a) 2, infinite, (b) 10, 0, (c) 0, 2, (d) 2, 0, (e) –4, 0,(f) 0, 0.5, (g) 0, 4,
(h) 10, 0, (i) 0, 0.2
5 (a) 18.10 V, (b) 7.36 V
6 9.96×10^4 Pa
7 (a) 0, (b) $0.86E/R$ A
8 (a) 0, (b) $0.86E$
9 0.95 μC
10 (a) 0, (b) 1.57 A, (c) 2.53 A, (d) 3.11 A, (e) 4 A
11 5.13 V
12 5637 Ω
13 (a) 1.26 A, (b) 1.72 A, (c) 1.90 A
14 0.03 g
15 (a) 8.61×10^4 Pa, (b) 7.41×10^4 Pa
16 (a) 198.0 V, (b) 190.2 V
17 (a) $-E$, (b) $-0.61E$
18 (a) 0, (b) $0.63E$

19 (a) 200°C, (b) 134.1°C, (c) 0°C
20 + gives growth, – decay.
21 (a) e^{8t}, (b) e^{-2t}, (c) e^{-12t}, (d) $1 + 2e^{2t} + e^{4t}$, (e) e^{-3t}, (f) e^{-2t}, (g) e^{-8t}, (h) $5e^{3t}$, (i) $0.4e^{4t}$

1.7

1 (a) $4 \lg x$, (b) $5 \lg x - \lg 2$
2 (a) $\lg b + 0.5 \lg 2 - \lg a - \lg c$, (b) $3 \lg a + 3 \lg b - 1.5 \lg c$
3 (a) 5.19, (b) –0.593, (c) 0.419
4 $v = 10.2\, e^{-0.1t}$
5 $\theta = 800\, e^{-0.2t}$
6 $Q = 2.6h^{2.5}$
7 $A = 400\, e^{-0.02t}$
8 $T = 50\, e^{0.3\theta}$
9 $I = 2430$ mA, $T = -51.3$
10 60.62×10^3

1.8

1 (a) 3.627, (b) 74.210, (c) 0.964, (d) –3.627, (e) 0.525, (f) 0.748
2 1.622 m
3 (a) $\sqrt{\left[\left(\frac{g\lambda}{2\pi} + \frac{2\pi\gamma}{\rho\lambda}\right)\frac{2\pi h}{\lambda}\right]}$, note when surface tension neglected $\sqrt{gh}$, (b) $\sqrt{\left(\frac{g\lambda}{2\pi} + \frac{2\pi\gamma}{\rho\lambda}\right)}$

Chapter 2

2.1

1 (a) 7.43 m/s, N 73° W, (b) 3.58 m/s, N 54° E, (c) 8.16 m/s, N 75° E
2 (a) 13 m, N 67° E, (b) 13 m, N 67° W, (c) 13 m, S 67° E, (d) 24.5 m, N 78° E
3 (a) $\overrightarrow{AC}$, (b) $\overrightarrow{BD}$, (c) $\overrightarrow{DB}$
4 (a) $\mathbf{b} - \mathbf{a}$, (b) $\mathbf{a} + \mathbf{b}$, (c) $\mathbf{a} - 3\mathbf{b}$, (d) $2\mathbf{b}$
5 7.8 N at 54° to AB
6 (a) $\overrightarrow{AD}$, (b) $\overrightarrow{AE}$, (c) 0, (d) $\overrightarrow{AC}$
7 1.36 N, 8.82 N
8 (a) 3.6, 56.3°, (b) 5.4, 21.8°, (c) 4.2, 45°
9 (a) $4\mathbf{i} + 6\mathbf{j}$, (b) $-8\mathbf{i}$, (c) $10\mathbf{i} + 9\mathbf{j}$
10 (a) $7\mathbf{i} + 5\mathbf{j}$, (b) $3\mathbf{i} - 1\mathbf{j}$, (c) $1\mathbf{i} - 4\mathbf{j}$
11 (a) $9\mathbf{i} + 2\mathbf{j}$, (b) $3\mathbf{i} + 4\mathbf{j}$, (c) $-13\mathbf{i} - 3\mathbf{j}$
12 (a) $\sqrt{74}$; $\frac{3}{\sqrt{74}}, \frac{7}{\sqrt{74}}, \frac{-4}{\sqrt{74}}$
(b) $\sqrt{38}$; $\frac{2}{\sqrt{38}}, \frac{3}{\sqrt{38}}, \frac{5}{\sqrt{38}}$,
(c) $\sqrt{38}$; $\frac{-3}{\sqrt{38}}, \frac{5}{\sqrt{38}}, \frac{2}{\sqrt{38}}$

13 $\sqrt{78}$; $\frac{2}{\sqrt{78}}$, $\frac{-5}{\sqrt{78}}$, $\frac{2}{\sqrt{78}}$
14 $17\mathbf{i} + 11\mathbf{j} + 5\mathbf{k}$, $\sqrt{255}$, $\sqrt{93}$, $\sqrt{2}$
15 112.6°
16 48.2°, 131.8°, 70.5°

2.3

1 (a) –j, (b) 1, (c) –j, (d) –1
2 (a) ±j4, (b) –2 ± j2, (c) 0.5 ± j1.1
3 (a) 4.12∠166°, (b) 5∠233°, (c) 3∠0°, (d) 6∠270°, (e) 1.4∠45°, (f) 3.61∠326°
4 (a) –2.5 + j4.3, (b) 7.07 + j7.07, (c) –6, (d) 0.68 + j2.72, (e) 1.73 + j1, (f) 1.5 – j2.6
5 (a) 1 + j6, (b) 5 – j2, (c) –14 + j8, (d) 0.23 – j0.15, (e) 0.1 – j0.8
6 (a) 5 – j2, (b) –2 – j1, (c) –1 + j7, (d) 1, (e) 12 + j8, (f) –10 + j6, (g) 11 – j2, (h) 12, (i) 10 + j5, (j) 0.9 + j1.2, (k) 0.23 – j0.15, (l) j1, (m) –0.3 + j1.1
7 (a) 20∠60°, (b) 50∠80°, (c) 0.1∠(–20°), (d) 0.5∠(–40°), (e) 5∠(–20°), (f) 0.4∠(–20°)
8 (a) 10∠(–30°), 8.66 – j5, (b) 10∠150°, –8.66 + j5, (c) 22∠45°, 15.6 + j15.6
9 (a) 5.5 + j2.6, 6.1∠25.3°, (b) –2 + j7, 7.3∠105.9°, (c) 3.7 + j4.5, 5.8∠50.6°
10 (a) 25∠90°, (b) 20∠75°, (c) 44.5∠83.3°, (d) 4∠(–30°), (e) 1.25∠15°, (f) 0.164∠9.2°
11 (a) 20 + j17.32 = 26.46∠40.9° V, (b) 26.46 sin(ωt + 40.9°) V
12 25∠(–30°) Ω
13 2∠(–36.8°) A
14 (a) 12 – j5 Ω, (b) 136.6 + j136.6 Ω, (c) 32.1 + j7.4 Ω, (d) 1.88 – j6.34 Ω, (e) 0.384 – j1.922 Ω, (f) j13.3 Ω
15 (a) 5 + j2 Ω, (b) 50 – j10 Ω, (c) 2 + j1 Ω, (d) 1.92 – j0.38 Ω, (e) –j125 Ω

Chapter 3

3.1

1 You might like to consider it to be like a swinging chain which, in itself, is rather like a form of simple pendulum.
2 See Figure S.2, $m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = F$
3 (a) $k_1x = M\frac{d^2y}{dt^2} + c\frac{dy}{dt} + (k_1 + k_2)y$,
(b) $T = I\frac{d^2\theta}{dt^2} + c\frac{d\theta}{dt} + k\theta$,
(c) $v = L\frac{di}{dt} + Ri$,
(d) $p = c\frac{d\theta}{dt} + cpq\theta$, where $\theta = \theta_o - \theta_i$

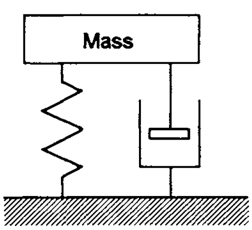


Figure S.2

3.2

1 (a) $E = 0.45L + 5$, (b) $R = 4.2L$
2 (a) T against $\sqrt{L}$, gradient $2\pi/\sqrt{g}$, intercept 0,
(b) s/t against t, gradient $a/2$, intercept u,
(c) e/θ against θ, gradient b, intercept a,
(d) R against θ, gradient $R_0\alpha$, intercept R_0,
3 $R = \dfrac{1200}{V} + 50$
4 $R = \dfrac{0.16}{d^2}$
5 $V = 16p^{-1}$
6 $T = 500p^{0.28}$
7 $C = 0.001n^3 + 30$
8 $s = 0.1v^2 + 0.5v$
9 $I = 0.001V^2$
10 $v = 2090\ e^{-t/12}$

Chapter 4

4.1

1 (a) $5x^4$, (b) $-8x^{-5}$, (c) $-6x$, (d) ½, (e) $4\pi x$, (f) $3\sec^2 3x$, (g) $-10\sin 2x$, (h) $8\ e^{x/2}$, (i) $-4\ e^{-2x}$, (j) $9\ e^{3x}$, (k) $-(5/6)x^{-3/2}$, (l) $-(12/3)x^{-3}$, (m) $-\dfrac{7}{2\sqrt{3}}x^{-3/2}$, (n) $-\dfrac{15}{8}x^{-4}$, (o) $\dfrac{\sqrt{3}}{2}x^{3/2}$, (p) $-24x^2 + 4x + 15$, (q) $5x\cos x + 5\sin x$, (r) $e^{x/2} + \frac{1}{2}x\ e^{x/2}$, (s) $(x^2 + 1)\cos x + 2x\sin x$, (t) $\dfrac{-13}{(x-6)^2}$, (u) $\dfrac{x-1}{2x^{3/2}}$, (v) $\dfrac{x\cos x - \sin x}{x^2}$, (w) $\dfrac{2\,e^{2x}(x^2 - x + 1)}{(x^2+1)^2}$, (x) $\dfrac{2\cosh 2x - \cosh 3x - 3\sinh 2x\sinh 3x}{(\cosh 3x)^2}$, (y) $\dfrac{7}{(2-7x)^2}$, (z) $\dfrac{1}{2(3-x)^{3/2}}$
2 (a) 2, (b) $-4\cos 2x$, (c) $6/x^4$, (d) $36x^2 - 2 - 2/x^3$, (e) $12x^2 + 12x$, (f) $\dfrac{3\sqrt{x} - \sqrt{x^3}}{4x^3}$
3 7 m/s, -4 m/s^2
4 $6\cos 2t - 9\sin 3t$ m/s, $-12\sin 2t - 27\cos 3t$ m/s^2
5 $0.03\cos 5t$ A/s
6 $50\ e^{-100t}$ V/s
7 πr^2
8 $4\pi r^2$
9 $-\dfrac{f_s}{c(1 + v/c)^2}$
10 $L_0(a + 2bT)$
11 (a) (2, −1) min., (b) (1, 7) max., (3, 3) min., (c) (1, −4) min., (−1, 4) max., (d) ($\pi/2$, 1) max., ($3\pi/2$, −1) min., (e) (−2, 23) max., (1, −4) min.
12 $r = h = 1$ m
13 $h = r = 4$ cm
14 ½L
15 0, 0.32L
16 47.7 V

17 3.33 m from smaller source
18 10 mA/s
19 $0.58L$
20 $-a/2b$
21 45°
22 $x = y$
23 $\frac{1}{2}(a - b)$
24 $r = 4/3$ m
25 6×6 cm
26 $4 \times 4 \times 2$ m

4.2

1 (a) $4x + C$, (b) $\frac{1}{2}x^4 + C$, (c) $\frac{1}{2}x^4 + \frac{5}{2}x^2 + C$,
(d) $\frac{3}{5}x^{5/3} - 2x^{3/2} + C$, (e) $4x + \frac{1}{5}\cos 5x + C$, (f) $-\frac{2}{3}\,e^{-3x} + C$,
(g) $8\,e^{x/2} + \frac{1}{3}x^3 + 2x + C$, (h) $4 \ln|x| + C$
2 (a) 15, (b) 8, (c) 1.10, (d) −2.25, (e) −4.5, (f) 2.67, (g) 0.83
3 (a) 116/3, (b) 11/3
4 ½
5 1/12
6 1
7 4½
8 (a) Diverges, (b) 1/18, (c) 1/3, (d) diverges
9 (a) $\frac{1}{40}(4x^2 - 1)(x^2 + 1)^4 + C$, (b) $\frac{1}{2}\,e^{4x-1} + C$,
(c) $\frac{1}{3}(x + 2)\sqrt{2x + 1} + C$, (d) $-\frac{1}{2}\cos x + C$,
(e) $2\sinh^{-1}\frac{x}{2} + \frac{1}{2}x\sqrt{x^2 + 4} + C$, (f) $\frac{2}{15}(3x + 2)(x - 1)^{3/2} + C$,
(g) $\ln\left|\dfrac{1 + \tan\frac{1}{2}x}{1 - \tan\frac{1}{2}x}\right| + C$, (h) $\sin^{-1}\frac{x}{3} + C$,
(i) $\frac{1}{6}\sin^3 2x - \frac{1}{10}\sin^5 2x + C$
10 (a) $\frac{2}{5}(x + 2)^{5/2} - \frac{4}{3}(x + 2)^{3/2} + C$, (b) $\dfrac{x}{\sqrt{x^2 + 1}} + C$,
(c) $\frac{1}{3}\cos^3 x - \cos x + C$, (d) $\frac{1}{2}\sin^{-1} 2x + C$,
(e) $\frac{1}{3}(x^2 + 2)^{3/2} + C$, (f) $\frac{1}{10}\tan^{-1}\frac{5x}{2} + C$,
(g) $\frac{1}{5}\sin^5 x - \frac{1}{7}\sin^7 x + C$, (h) $\frac{1}{4}\tan^4 x + C$,
(i) $\frac{2}{3}\tan^{-1}\left(\frac{1}{3}\tan\frac{x}{2}\right) + C$
11 (a) ln 2, (b) 1/90, (c) $\pi/4$, (d) $\pi/4$, (e) $\pi/8$
12 (a) $\frac{1}{3}x^3 \ln|x| - \frac{1}{9}x^3 C$, (b) $\frac{1}{2}x\,e^{2x} - \frac{1}{4}\,e^{2x} + C$,
(c) $x^3 \sin x + 3x^2 \cos x - 6x \sin x - 6\cos x + C$,
(d) $\frac{1}{25}\sin 5x - \frac{1}{5}x\cos 5x + C$, (e) $\frac{1}{2}x^2 \ln|3x| - \frac{1}{4}x^2 + C$,
(f) $\frac{1}{2}x - \frac{1}{4}\sin 2x + C$
13 (a) $\frac{\pi}{2} - 1$, (b) $\frac{1}{16}\pi^2 + \frac{1}{4}$, (c) 2
14 (a) $\frac{1}{4}x^2 + \frac{3}{4}x + \frac{9}{8}\ln|2x - 3| + C$, (b) $-\frac{1}{2}x - \frac{1}{4}\ln|1 - 2x| + C$,
(c) $\frac{1}{2}x + \frac{1}{5}\ln|x - 1| - \frac{9}{20}\ln|2x + 3| + C$,
(d) $\frac{1}{6}\ln|x - 1| + \frac{1}{2}\ln|x + 1| - \frac{1}{6}\ln|2x + 1| + C$,
(e) $-\frac{1}{4}\ln|x| + \frac{3}{8}\ln|x - 2| - \frac{1}{8}\ln|x + 2| + C$,
(f) $\frac{2}{3}\ln|x - 1| + \frac{5}{6}\ln|2x + 1| + C$,
(g) $\frac{1}{2}x^2 + 2x + \frac{2}{3}\ln|x - 1| + \frac{5}{6}\ln|2x + 1| + C$,
(h) $\ln|x - 3| - \ln|x - 2| + C$, (i) $6\ln|x| - \ln|x + 1| - \dfrac{9}{x + 1} + C$,

(j) $2\ln|x| - 2\ln|x-1| + \ln|x^2+4| + 2\tan^{-1}\frac{1}{2}x + C$,
(k) $-\frac{1}{x} - \tan^{-1}x + C$
15 (a) $Mh^2/18$, (b) $Mh^2/6$
16 $5Mr^2/2$
17 (a) $\frac{1}{12}ML^2$, (b) $M\left(\frac{1}{12}L^2 + d^2\right)$
18 (a) 1, (b) 7, (c) 16, (d) 0.5
19 $2A/\pi$
20 $0.623N_0$
21 (a) 4.92, (b) 1.15, (c) 1.23, (d) 0.707, (e) 1.35
22 $V/2$

Chapter 5

5.1

1 (a) $m\frac{dv}{dt} + kv^2 = mg$, (b) $m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = 0$,
(c) $m\frac{dv}{dt} + kv = 0$, (d) $m\frac{dv}{dt} + kv^2 = mg$, (e) $\frac{dI}{dx} + kx = 0$,
(f) $A\frac{dh}{dt} + \sqrt{2gh} = 0$, (g) $\frac{A}{\rho g}\frac{dp}{dt} + \sqrt{\frac{2p}{\rho}} = q_1$
2 As given in the problem
3 (a) $y = x\,e^x$, (b) $y = \frac{1}{\omega}\sin\omega t + 2\cos\omega t$, (c) $y = (2 + x^2)\,e^{-x}$
4 $y = -\frac{w}{2EI}\left(\frac{1}{2}L^2x^2 - \frac{1}{3}Lx^3 + \frac{1}{12}x^4\right)$

5.2

1 (a) $y = \ln x + A$, (b) $y = 2\sin\frac{1}{2}x + A$, (c) $-1/y = x + A$,
(d) $\ln y = -2x + A$ or $y = C\,e^{-2x}$, (e) $\tan^{-1}y = x^2 + A$,
(f) $y = \ln(x^3 + A)$ or $e^y = x^3 + A$, (g) $y = 4x + x^3 + A$,
(h) $-1/y = 2x + A$, (i) $e^y = x^3 + A$, (j) $y^2 = x + A$,
(k) $y = A\,e^{-1/x}$, (l) $y = -2/(x^2 + A)$
2 $y = 1/(2 - x^2)$
3 $V = V_0\,e^{-t/RC}$
4 $N = N_0\,e^{-kt}$
5 $i = \frac{V}{R}(1 - e^{-Rt/L})$
6 $y = 2 - e^{-x}$
7 $v = 10 - 10\,e^{-t}$
8 $T = T_0\,e^{\mu\theta}$
9 3.41 hours
10 RC
11 51.4°C
12 As given in the problem
13 10.9 V
14 $i = I(1 - e^{-Rt/L})$
15 (a) $x = 1 - e^{-t/2}$, (b) $x = 8\,e^{-t/2} + 4t - 8$
16 $x = 5(1 - e^{-t/\tau})$
17 (a) 36.8%, (b) 13.5%
18 12 s
19 (a) $2\frac{dx}{dt} + x = 45$, (b) $x = 45 - 25\,e^{-t/2}$

5.3

1 (a) $y=\frac{3}{2}x^2+4x+4$, (b) $y=\frac{3}{5}\,e^{2x}+\frac{2}{5}\,e^{-3x}$, (c) $y=e^{-2x}+2x\,e^{-2x}$, (d) $y=\frac{5}{2}\,e^{x}-\frac{5}{2}\,e^{-x}$, (e) $y=2\,e^{x}-e^{2x}$, (f) $y=2\,e^{-x}+x\,e^{-x}$

2 (a) $y=A\,e^{x/2}+B\,e^{-x}$, (b) $y=(A+Bx)\,e^{3x}$, (c) $y=(A+Bx)\,e^{5x}$, (d) $y=e^{2x}(A\cos x+B\sin x)$, (e) $y=e^{x}(A\cos 2x+B\sin 2x)$, (f) $y=A\,e^{2x}+B\,e^{-5x}$, (g) $y=e^{2x}(A\cos x+B\sin x)$, (h) $y=A\,e^{2x}+B\,e^{-4x}$,

3 (a) $y=-\frac{3}{4}\,e^{-5x}+\frac{3}{4}\,e^{-x}$, (b) $y=e^{3x}(3\cos 4x-2\sin 4x)$, (c) $y=(2+4x)\,e^{-3x/2}$, (d) $y=2\,e^{3x}-e^{4x}$, (e) $y=(1+2x)\,e^{-x}$, (f) $y=e^{3x}(\cos 4x-\sin 4x)$

4 (a) $y=A\,e^{2x}+B\,e^{-2x}+\frac{2}{5}\,e^{3x}$, (b) $y=e^{-3x/2}\left(\cos\sqrt{\frac{7}{2}}\,x+\sin\sqrt{\frac{7}{2}}\,x\right)+\frac{3}{4}x-\frac{1}{16}$, (c) $y=A\,e^{4x}+B\,e^{2x}+\frac{21}{85}\cos x-\frac{18}{85}\sin x$, (d) $y=A\,e^{2x}+B\,e^{3x}+\frac{1}{2}e^{x}$, (e) $y=A\,e^{2x}+B\,e^{3x}+e^{x}-\frac{1}{4}\,e^{-x}$, (f) $y=A\,e^{2x}+B\,e^{-x}+\frac{1}{4}\cos 2x-\frac{3}{4}\sin 2x$, (g) $y=A\,e^{x}+B\,e^{2x}-2-3x-x^2$,

5 $y=e^{3x}-3\,e^{2x}$

6 $y=\cos 3x-\frac{2}{15}\sin 3x+\frac{1}{5}\sin 2x$

7 $x=A\,e^{-4t}+B\,e^{-t}$

8 (a) $x=0.2\cos 3t$, (b) $x=e^{-t}(0.2\cos 2.83t+0.070\sin 2.83t)$

9 0.44

10 (a) 5 rad/s, (b) 1.25

11 316 rad/s, 6.3 N s/m

12 Over damped

13 6 N s/m

14 2.6 rad/s, 0.76

15 $e^{-t}(0.2\cos 2.24t+0.22\sin 2.24t)$

16 $\theta=\frac{4}{9}\pi\,e^{-t}-\frac{1}{9}\pi\,e^{-4t}$

17 As given in the problem

Chapter 6

6.1

1 (a) $\frac{2}{s^3}$, (b) $\frac{6}{s^4}$, (c) $\frac{a}{s^2-a^2}$

2 (a) $\frac{4}{s}$, (b) $\frac{3}{s^2}-\frac{1}{s}$, (c) $\frac{1}{s-3}$, (d) $\frac{2}{s^2}+\frac{3}{s-1}$, (e) $\frac{2}{s^3}+\frac{4}{s+2}$, (f) $\frac{2}{s^3}+\frac{2}{s^2}+\frac{1}{s}$, (g) $\frac{6}{s^2+9}$, (h) $\frac{15}{s^2-9}$, (i) $\frac{3}{s^2+36}$, (j) $\frac{1}{(s+3)^2}$, (k) $\frac{4}{s}-\frac{6}{s^2+9}+\frac{1}{s-2}$, (l) $\frac{6}{(s+2)^2}$, (m) $\frac{2}{s^2-1}$, (n) $\frac{s-3}{s^2-6s+10}$, (o) $\frac{s^2+4s+5}{(s+1)^3}$, (p) $\frac{4}{(s+1)(s^2+2s+5)}$, (q) $\frac{s^2+9}{(s^2-9)^2}$, (r) $\frac{2s(s^2+27)}{(s^2-9)^3}$, (s) $\frac{6}{(s+3)^4}$

3 (a) $\frac{2}{s^3}+\frac{3}{s^2}+\frac{2}{s}$, (b) $\frac{2}{s}+\frac{12}{s^2+9}$, (c) $\frac{1}{s-4}+\frac{s}{s^2-4}$, (d) $\frac{2}{s}+\frac{5}{s-3}$, (e) $\frac{s}{s^2+4}+\frac{s}{s^2+9}$, (f) $\frac{6}{s^4}+\frac{4}{s+1}$

4 (a) $\frac{2}{(s+3)^2+4}$, (b) $\frac{2}{(s-4)^3}$, (c) $\frac{s-2}{(s-2)^2+1}$
5 (a) $e^{-5s}\frac{1}{s}$, (b) $1\,e^{-4s}$ (c) $\frac{3\,e^{-10s}}{s^2}$
6 $\frac{3(1-e^{-s})}{s(1+e^{-s})}$
7 $\frac{1}{s^2}-\frac{e^{-s}}{s(1-e^{-s})}$
8 (a) $\frac{1}{s(1+e^{-s})}$, (b) $\frac{1}{s^2(1+e^{-s})}-\frac{e^{-s}}{s(1-e^{-2s})}$, (c) $\frac{1-e^{-s}}{s^2(1+e^{-s})}$
9 (a) e^{2t}, (b) 5, (c) $\cos 4t$, (d) $\sinh 3t$, (e) $e^{2t}\sin 5t$,
(f) $\frac{1}{6}t^3\,e^{-3t}$, (g) $(t-2)u(t-2)$, (h) $(t-3)\,e^{-2(t-3)}u(t-3)$
10 (a) $e^{-2t}+2\,e^{3t}$, (b) $2\,e^{t}+e^{-2t}$, (c) $\cos 2t-e^{-t}$, (d) $e^{-2t}+2t\,e^{-2t}$
11 (a) 0, (b) 1
12 (a) 2, (b) 0

6.2

1 (a) $x=3\,e^{2t}-3\,e^{-t/2}$, (b) $x=\frac{2}{5}-\frac{2}{5}e^{-5t}$, (c) $x=\frac{1}{4}-\frac{1}{4}\cos 2t$,
(d) $x=e^{-t}-e^{-t}\cos t$, (e) $x=\frac{1}{2}\,e^{t}-\frac{1}{2}t\,e^{t}+\frac{1}{2}\cos t$,
(f) $x=2\,e^{t}-1$, (g) $x=\frac{4}{17}\cos t+\frac{2}{17}\sin t-\frac{4}{17}\,e^{-4t}$,
(h) $x=3-3\cos t+\sin t$,
(i) $x=\frac{1}{5}\,e^{-3t}+t\,e^{2t}-\frac{1}{5}\,e^{2t}$,
(j) $x=\frac{3}{65}\,e^{3t}+\frac{1}{20}\,e^{-2t}-\frac{1}{52}\sin 2t-\frac{5}{52}\cos 2t$,
(k) $x=\frac{5}{36}-\frac{1}{6}t-\frac{1}{12}\,e^{3t}+\frac{121}{252}\,e^{6t}+\frac{369}{252}\,e^{-t}$

6.3

1 $3/s$
2 2
3 $5/s^2$
4 e^{-5t} V
5 (a) $(5/3)(1-e^{-3t})$ V, (b) $5e^{-3t}$ V
6 (a) $6(1-e^{-t})$ V, (b) $6e^{-t}$ V
7 (a) $2(1-e^{-2t})$ V, (b) $\frac{1}{2}[t-\frac{1}{2}(1-e^{-2t})]$ V
8 (a) $5/(s-1)-4/(s-2)$, (b) $4/(s+1)-3/(s+2)$,
(c) $2/(s+1)-3/(s+1)^2$
9 $24-12e^{-2t}-4e^{-4t}$
10 $8e^{-2t}-8t\,e^{-2t}$
11 (a) Critical, (b) overdamped, (c) underdamped
12 $-1.5+3.0t+1.5e^{-2t}$
13 $0.5-e^{-t}+0.5e^{-2t}$
14 $0.5(e^{-t}-e^{-3t})$
15 $0.5(1-e^{-10t})$
16 10, 0.05
17 Underdamped
18 Critically damped
19 1/53 s
20 (a) $0.01s\ \Omega$ in series with -0.002 V, in parallel with $0.2/s$ A,
(b) $0.5/s$ MΩ in series with $5/s$ V, in parallel with 10 μA.
21 (a) $10+0.002s\ \Omega$, (b) $\frac{0.02s}{10+0.002s}\ \Omega$
22 $i(t)=\frac{V_0}{R}\,e^{-t/RC}$

23 $i(t) = \frac{V}{R} - \frac{V}{R}\,e^{-Rt/L}$

24 $i(t) = \frac{1}{R}\left(\delta(t) - \frac{1}{RC}\,e^{-t/RC}\right)$

Chapter 7

7.1

1 0, 1, 0, −1, 0, 1, ...
2 (a) 0, 1, 2, 3, 4, (b) 1, 0.37, 0.13, 0.05, 0.007
3 (a) 116, (b) 0.75
4 (a) $(-1)^{k-1}$, (b) $5k$, (c) $2.5 - 0.5k$
5 (a) 0, 1, 4, 9, 16, (b) 1, 2.72, 7.39, 20.09, 54.60, (c) 0, 2.5, 6, 10.5, 16
6 13
7 0.5
8 (a) 0.25^k, (b) $2(-1)^k$, (c) $3 + 0.1^k$
9 (a) 0.1, 0.01, 0.001, (b) 5.1, 5.01, 5.001, (c) −1, +1, −1
10 (a) 222, (b) 9.998, (c) 28.70
11 (a) 7.5, (b) 23.98, (c) 1023
12 (a) 12, (b) 16, (c) 48
13 (a) Convergent, (b) divergent
14 (a) Convergent, (b) divergent
15 (a) Divergent, (b) convergent, (c) convergent, (d) convergent, (e) convergent, (f) divergent
16 (a) $1 + 4x + 6x^2 + 4x^3 + x^4$, (b) $1 + \frac{3x}{2} + \frac{3x^2}{8} - \frac{x^3}{16}$, (c) $1 + \frac{5x}{2} + \frac{35x^2}{8} + \frac{105x^3}{16}$ (d) $1 - 0.25 + 0.062 - 0.015$, (e) $2 + \frac{x}{4} - \frac{x^2}{64} + \frac{x^3}{512}$
17 (a) $1 + 12x + 66x^2 + 220x^3 + \dots$, (b) $1 + 4x + 12x^2 + 32x^3 + \dots$, (c) $3^{2/5}\left(1 - \frac{4}{15}x - \frac{4}{75}x^2 - \frac{64}{3375}x^3 + \dots\right)$, (d) $1 + x + x^2 + x^3 + \dots$, (e) $1 - \frac{3}{2}x + \frac{27}{8}x^2 - \frac{135}{16}x^3 + \dots$, (f) $1 + 2x^3 + 3x^6 + 4x^9 + \dots$
18 1.013 2
19 As given in the problem
20 $1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \dots$
21 $x + \frac{1}{3}x^3 + \frac{2}{5}x^5 + \dots$
22 (a) $1 + 2x + 2x^2 + \frac{4}{3}x^3 + \dots$ (b) $1 + x - \frac{1}{2}x^3 + \dots$, (c) $1 - \frac{1}{2}x + \frac{1}{4}x^2 - \frac{5}{8}x^3 + \dots$, (d) $x + \frac{x^2}{2!} + \frac{2x^3}{3!} + \dots$, (e) $1 + \frac{1}{2}x^2 + \frac{5}{24}x^4 + \dots$, (f) $1 - x^2 + \frac{1}{3}x^4 - \dots$
23 As given in the problem
24 As given in the problem
25 As given in the problem
26 Reduced by 6%
27 Increased by 1%

7.2

1 Second and fourth (and a d.c. term)
2 $\frac{4A}{\pi}\left(\sin \omega t + \frac{1}{3}\sin 3\omega t + \frac{1}{5}\sin 5\omega t + \dots\right)$

3 $\frac{A}{2}-\frac{4A}{\pi^2}\left(\sin\omega t+\frac{1}{9}\sin 3\omega t+\frac{1}{25}\sin 5\omega t+\ldots\right)$
4 $\frac{2A}{\pi}-\frac{4A}{\pi}\left(\frac{1}{3}\cos 2\omega t+\frac{1}{15}\cos 4\omega t+\frac{1}{35}\cos 6\omega t+\ldots\right)$
5 (a) Only even harmonics, (b) only odd harmonics
6 $\frac{4A}{\pi}\left(\cos\omega t-\frac{1}{3}\cos 3\omega t+\frac{1}{5}\cos 5\omega t-\ldots\right)$
7 (a) Odd sines, (b) a_0, even sines and cosines
8 (a) 10/6, (b) 0
9 (a), (b), (d) odd, (c), (e) even
10 (a) sine, (b) a_0 and cosine, (c) a_0 and cosine
11 $2\left(\sin t-\frac{1}{2}\sin 2t+\frac{1}{3}\sin 3t+\ldots\right)$
12 (a) 0.5, 0.31, 0.16, 0.11, 180°, 180°, 180°,
(b) 1.57, 1.19, 0.5, 0.13, −32°, 180°, −12°
13 $0.32 + 0.5\cos 100t + 0.21\cos 200t$ mA
14 $0.5\cos(100t - 90°) + 0.21\cos(200t - 90°)$ A
15 $3.2\sin(100t + 90°) + 3.2\sin(200t + 90°)$ mA

Chapter 8

8.2

1 (a) 1, (b) 1, (c) 0
2 (a) $a \cdot b$, (b) $a + b + c$, (c) b, (d) $a \cdot \overline{b}$, (e) $a + \overline{b}$

8.3

1 (a) $(b + c) \cdot d + a$, (b) $(a + \overline{a}) \cdot (b + \overline{b})$, (c) $a \cdot b + \overline{a} \cdot \overline{b}$,
(d) $a \cdot \overline{c} + \overline{a} \cdot b + \overline{a} \cdot \overline{b} \cdot \overline{c}$
2 (a)

a	b	c	$(a+\overline{b})+(a+\overline{c})$
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

(b)

a	b	$b \cdot a$	$\overline{b}$	$a \cdot b+\overline{b}$	$\overline{a}$	$\overline{a} \cdot (a \cdot b+\overline{b}) \cdot \overline{b}$
0	0	0	1	1	1	1
0	1	0	0	0	1	0
1	0	0	1	1	0	0
1	1	1	0	1	0	0

3 (a) $A \cdot B + C \cdot (D + E)$, (b) $A + (B + C)$, (c) $A + B + C$

4

a	b	$\overline{a}$	$\overline{b}$	$a\cdot\overline{b}$	$\overline{a}\cdot b$	$(a\cdot\overline{b})+(\overline{a}\cdot b)$
0	0	1	1	0	0	0
0	1	1	0	0	1	1
1	0	0	1	1	0	1
1	1	0	0	0	0	0

5 (a) See Figure S.3(a), (b) see Figure S.3(b), (c) see Figure S.3(c)

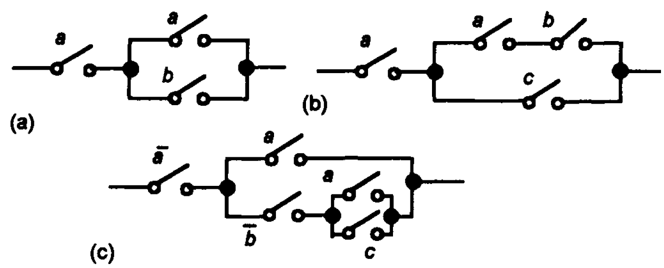


Figure S.3

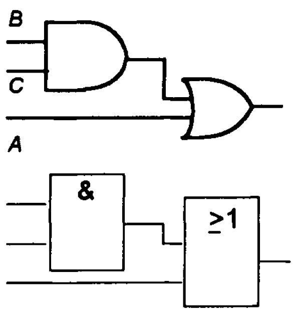


Figure S.4

6 (a) $(A+B)\cdot B\cdot C+A$, $B\cdot C+A$, see Figure S.4,
(b) $\overline{A\cdot B}\cdot B$, $\overline{A}\cdot B$, see Figure S.5

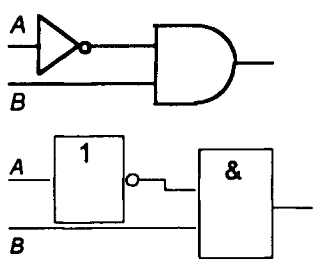


Figure S.5

7 (a) See Figure S.6(a), (b) see Figure S.6(b), (c) see Figure S.6(c), (d) see Figure S.6(d)

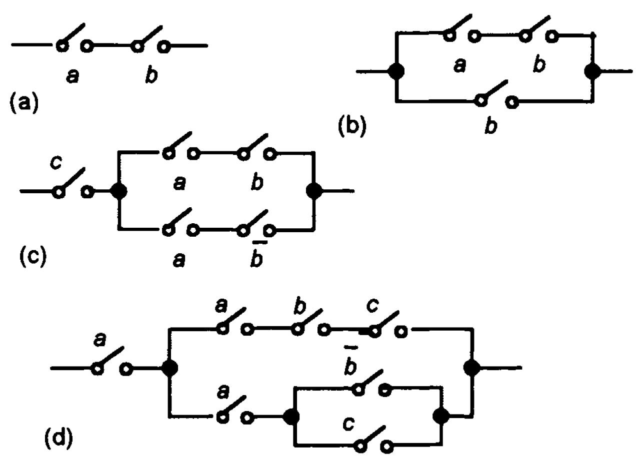


Figure S.6

8 (a) $\bar{a}\cdot\overline{b}\cdot c+a\cdot b\cdot\overline{c}$, (b) $c\cdot(a\cdot b+\overline{a}\cdot b)$

9 (a) $A\cdot B\cdot(A+B+C)$, $A\cdot B\cdot C$,
(b) $(A\cdot B+\overline{A}\cdot C)+(\overline{A}\cdot B+B\cdot C)$, $B+\overline{A}\cdot C$
See Figure S.7

10 (a) $\overline{A}\cdot\overline{B}+A\cdot B$, (b) $B\cdot\overline{C}$, (c) $\overline{A}\cdot\overline{C}+A\cdot B+A\cdot C$, o $\overline{A}\cdot\overline{C}+\overline{A}\cdot B+A\cdot C$, (d) $\overline{B}\cdot\overline{C}$, (e) C, (f) $C\cdot D$, (g) $A\cdot B+\overline{B}\cdot\overline{C}$

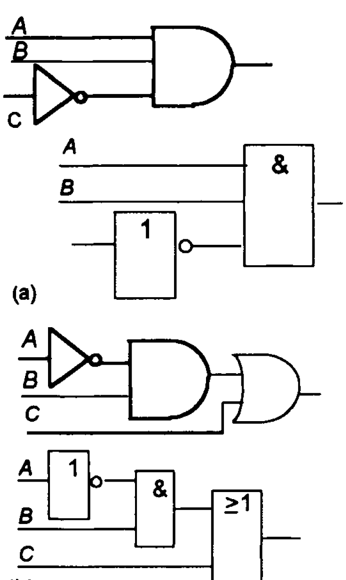


Figure S.7

Chapter 9

9.1

1 4/50
2 4/52
3 (a) 0.004, (b) 0.96
4 0.85
5 0.97
6 (a) 0.01, (b) 0.02, (c) 0.03, (d) 0.96
7 (a) 1/3, (b) 2/3
8 (a) 40 320, (b) 840, (c) 30
9 1260
10 (a) 21, (b) 21, (c) 1
11 (a) 77/92, (b) 11/69, (c) 1/276
12 (a) 1/15, (b) 7/15
13 (a) 1/15, (b) 8/15
14 1000
15 0.5

9.2

1 3.0
2 4/3
3 40, 31.6
4 5
5 0.237
6 0.24
7 (a) 0.02, 0.06, 0.22, 0.32, 0.28, 0.08, 0.02, (b) 69.3, 2.3
8 (a) 0.4, (b) 7
9 0.122, 0.270, 0.285, 0.190, 0.090, 0.032, 0.009, 0.002, 0.000 4, 0.000 1
10 0.132
11 (a) 0.817, (b) 0.016
12 0.366
13 0.005, 0.029, 0.078, 0.138, 0.181, 0.185, 0.154, 0.108, 0.064, 0.033
14 0.016
15 100, 9.9
16 0.001
17 0.135
18 0.297
19 20
20 (a) 6.68, (b) 17.75
21 (a) 9.68, (b) 11.5
22 (a) 0.091 3, (b) 0.091 3, (c) 0.817 4
23 50 ± 6.6 Ω
24 0.34
25 0.954
26 0.185 9
27 (a) 0.315 6, (b) 0.726 0

9.3

1. 2.134 mm, 0.011 mm
2. 0.05 A
3. 51.12 Ω, 0.08 Ω
4. 72
5. 4.9 ± 0.3 mm^2
6. 10 000
7. 39.0 kV, 0.11 kV
8. (a) 50 ± 3.6, (b) $10\,000 \pm 500$, (c) $1\,000\,000 \pm 52\,000$, (d) 2 ± 0.1
9. Diameter
10. (a) 150 ± 3.6, (b) 5000 ± 250, (c) 2 ± 0.1

Index